AI Use Case Assessment

Analyst: Gabriel Ortiz, AI Governance Lead
Organization: Alderwyn Health
System: CarePivot

Instructions

A use-case assessment is the structured evaluation of whether and how an AI use case should proceed. It is conducted before any data is touched or model trained. Each of the eight sections below opens with an objective that explains what the section is for and why it belongs in the assessment, followed by the guiding questions to answer. Work top to bottom.

Section 1: Business Problem & Context

Objective: Establish the real-world problem the use case is meant to solve. State the problem in business terms and avoid referencing AI. The goal is to produce a clear-eyed account of what is going wrong today, what the status quo costs, and why acting now is justified. This way, the case rests on a genuine need rather than technology enthusiasm. If the problem cannot be articulated without mentioning the model, that is a signal to pause.

What is going wrong today?

Patients leave our hospitals and come back within thirty days, and each return usually marks a failed handoff from hospital to home. The transitional care that prevents these returns exists and works, but the team can reach only about 15 percent of diabetic discharges. Today we choose that 15 percent with unit-nurse judgment and a length-of-stay rule, and everyone involved can name patients the method missed.

What does the status quo cost?

Four consecutive years of Medicare readmission penalties, roughly four to six million dollars annually, and the human cost behind the penalty: patients back in a hospital bed within weeks who might have been caught by a phone call. The current selection method also spends scarce nursing attention unevenly, because instinct varies by shift and unit while the rule measures only length of stay.

Why act now?

The board has directed measurable improvement within two penalty cycles, a prototype already exists, and the transitional care team itself is asking for better targeting. Acting now also means the workflow and fairness questions get decided deliberately, on the record, before any tool touches a discharge decision.

SECTION 2: DECISION(S) THE USE CASE SUPPORTS

Objective: Pin down the specific human decision or process the AI system is meant to inform. List who owns the decision and/or process. This section fixes the system's scope and posture: determining whether the system offers decision-support with a human in the loop or automates the decision outright. This section also sets the purpose limitation by drawing the red lines (i.e., the uses the output must never be put to). Everything downstream to the decision or process should inherit these boundaries.

What does the decision change?

Which diabetic inpatients the transitional care team enrolls at discharge, within a fixed capacity of roughly 15 percent of diabetic discharges. Because capacity is fixed, this is an allocation decision: every patient ranked in displaces a patient who would otherwise have been chosen. For enrolled patients the decision changes the support they receive in the most fragile weeks after discharge. For displaced patients it changes the same thing, invisibly.

Who is the human decision-maker?

Today, the transitional care director and her nurses, informed by unit-nurse referrals and the length-of-stay rule. Under this assessment's recommendation that does not change: the director and her team decide whom to enroll. CarePivot proposes an ordering. People do the choosing.

Is the AI system decision support, or will it make decisions automatically?

Recommendation: decision support as a ranked worklist for the transitional care director. The team retains full authority to depart from the list, and clinical judgment is never penalized for overriding a score. A discharge-navigator flag should be considered only after the tool has earned bedside trust and its error patterns are understood. Automatic enrollment should not be considered on current evidence: this is a zero-sum allocation, the model is unvalidated, and fairness cannot yet be measured for every group we serve.

What is the AI system NOT to be used for?

CarePivot output must not influence discharge timing, bed management, treatment decisions, insurance or billing matters, or any decision beyond transitional care enrollment. No patient may be removed from care they would otherwise receive because of a score. Scores must never be presented to patients or families as a judgment about them. Any new use requires a new assessment.

SECTION 3: AI / ML FRAMING

Objective: Translate the business problem into a precise technical problem type. Name the unit of analysis, the target being predicted, and the baseline the model must beat. Then, test honestly whether a simpler non-AI method would do the job. The objective is to keep the team disciplined about whether machine learning is actually warranted, and to make the model's added value (and added risk) an honest comparison rather than an assumption.

What is the technical problem being addressed?

Binary classification at the encounter level. Given what is known about a diabetic inpatient at discharge, estimate the probability of an unplanned return within thirty days. The output is a risk score used to rank discharges for transitional care attention.

What is the unit, target, or baseline?

The unit of analysis is a hospital stay, an encounter, which matters because our extract contains many patients with multiple stays. The target requires an explicit framing decision: the extract records the outcome three ways (no readmission, readmission after thirty days, readmission within thirty days), and this assessment adopts readmission within thirty days versus everything else, documented as a decision rather than a default. The baseline to beat is current practice, the length-of-stay rule plus nurse judgment. If CarePivot cannot select patients who benefit more than the current method's selections, it should not be deployed.

Could a simpler, non-AI/ML solution solve this problem?

Partially. Better referral checklists and more consistent unit-nurse processes would improve today's selection without any model, and those improvements are worth making regardless. What simpler methods cannot do is weigh dozens of factors consistently across every discharge on every shift. A learned model is a reasonable candidate for that

job, provided the comparison to current practice stays honest and the fairness questions in this file are answered first.

SECTION 4: AFFECTED PARTIES

Objective: Identify everyone the use case touches, not just its users. Identify and list the AI subjects (those being scored or acted on), the decision-makers, and the parties affected downstream of the system's deployment and decision-making. Pay particular attention to those affected who have no say. Surfacing protected groups here is what makes fairness testing possible later. This section is the foundation for the success criteria in §5 and the harm inventory in §7.

Who are the AI subjects?

Diabetic inpatients at the moment of discharge, scored at one of the most vulnerable points in their care. They are the people ranked by the system.

Who are the decision-makers?

The transitional care director and her team, as the deciding humans. The Chief of Medicine's clinicians, whose referrals and trust the tool depends on. The Chief AI Officer, as approver of record for program decisions.

Who is affected downstream?

Enrolled patients, who receive calls, medication reconciliation, and visits. Displaced patients, who would have been chosen under current practice and now are not, and who never learn a model was involved. The bedside nurses whose judgment the tool touches. The informatics team. Compliance. Medicare and the Office for Civil Rights. The mission neighborhoods whose care patterns this tool will quietly shape.

Which protected groups are present?

The extract records race, gender, and age. Race is missing for about 2 percent of encounters, and age arrives only in ten-year bands. Under Section 1557, discrimination on the basis of race, color, national origin, sex, age, or disability is prohibited in programs like ours, and the decision support rule expects us to identify and mitigate discrimination risk in tools like this one. Where race is unrecorded, fairness for those patients cannot be measured as the data stands. That is a program problem to fix, not a footnote.

Who is affected but has no say?

The displaced patient, above all: ranked out by a tool they never saw, appearing in no complaint queue and no dashboard unless we deliberately count them. This assessment recommends counting them from day one.

SECTION 5: DEFINITION OF SUCCESS

Objective: State what success means concretely across three lenses: the business outcome, the model metrics that fit the data (not a default like accuracy), and the fairness criterion with its named groups and threshold. Defining failure just as concretely is part of the job. The goal is a target that is measurable and testable rather than aspirational. With this, later stages can verify whether "success" was met.

What is the specific business outcome the solution will address?

Direct transitional care capacity to the diabetic discharges most likely to return within thirty days, so that the same 15 percent capacity prevents more readmissions than current practice, without directing help away from the neighborhoods our mission names.

What metrics will be measured to determine success?

Readmission outcomes among enrolled versus comparable non-enrolled patients. Selection quality against the incumbent method: of the patients each method would choose, how many actually returned. Override rates and patterns, as a signal of clinical trust and of what the model misses. Enrollment patterns by race, age band, sex, and neighborhood. Penalty exposure over time, acknowledged as a lagging, condition-mix-driven measure.

What fairness criterion will be used to identify whether the solution performs equitably?

Per-group comparison of the model's error rates and of enrollment rates across race, sex, age bands, and service neighborhoods, with the recorded-race gaps addressed first so the measurement is honest. The working criterion is that the tool should not select systematically away from any group at equal clinical risk. Precise measures and thresholds are finalized during evaluation and recorded before any reliance.

What thresholds will determine the AI system's success?

To be pre-registered before deployment and graded at retirement. Working targets: selection quality measurably better than the length-of-stay rule plus judgment. No material adverse per-group gap in enrollment or error rates without documented justification and mitigation. Override review running monthly with findings fed back. The displaced-patient picture reviewed by group and neighborhood from day one.

What does failure look like?

A tool that directs help away from the mission neighborhoods would be failure regardless of the penalty number. So would selection quality no better than the current rule, clinician trust so low the list is ignored, or a fairness audit we cannot run because we never fixed the data. Failure is defined now so nobody redefines it later.

SECTION 6: CONSTRAINTS

Objective: Catalogue the boundaries the use case must operate within before design begins. Naming the legal/regulatory, ethical/fairness, data/privacy, operational, and resource constraints up front prevents late-stage surprises and defines the guardrails the build must respect. Each constraint named here should have a visible home in the design and, where relevant, in the compliance record.

What legal and regulatory considerations inform the AI system's design?

Section 1557 of the Affordable Care Act (42 U.S.C. 18116) and the patient care decision support tool provision (45 CFR 92.210), which apply directly to a tool of this kind. HIPAA's Privacy and Security Rules over the encounter data. The Medicare readmission program as financial context, with its measured-conditions accuracy note. The clinical decision support classification question under 21 U.S.C. 360j(o), which this program should analyze and document rather than assume.

What ethical and fairness principles inform the AI system's design?

The bedrock principle that the tool informs care and never rations it: enrollment decisions remain human, clinical override is protected, and the displaced are counted. Fairness measured rather than assumed, with the Obermeyer mechanism treated as the default failure mode to rule out, given our utilization-heavy features.

What data protection and privacy considerations inform the AI system's design?

The extract is protected health information. Access is limited to the program team under the system's HIPAA controls, use for this development should be confirmed against our privacy practices, and no patient data leaves Alderwyn's environment. Any future data corrections, including race recording, must follow privacy-compliant processes.

What operational considerations inform the AI system's design?

The tool must fit the transitional care team's actual workflow and produce its ranking by discharge time, not after. It must fail safe: if CarePivot is unavailable, selection continues under current practice. Nurse training, a documented override path, and monthly override review are required. Alert and list fatigue are design constraints, not afterthoughts.

What resource considerations inform the AI system's design?

A four-person informatics team, a governance office newer than the program it governs, and a transitional care team already at capacity. The program must not add documentation burden to bedside nurses. A tool the team stops opening by month two is a failed tool, whatever its statistics say.

SECTION 7: INITIAL RISK & HARM INVENTORY

Objective: Produce a first-pass catalogue of the positive and negative harms the use case could cause, including who bears each and the worst/best plausible case. This is not the full risk assessment that comes later. The objective is to see the hazards clearly enough in order to judge whether and how to proceed, and to hand the downstream risk assessment a starting list rather than a blank page.

What negative harms could the AI system cause?

Selection that systematically directs help away from patients with historically less access, through utilization-shaped proxies, the Obermeyer pattern reproduced locally. Unmeasurable fairness for patients whose race is unrecorded. Flattered performance estimates if the same patient's stays appear on both sides of an evaluation. Erosion of clinician trust if the list contradicts bedside judgment without explanation. Quiet harm to displaced patients that no complaint process will ever surface.

What positive harms could the AI system cause?

More readmissions prevented with the same capacity. Earlier, better-targeted support in the fragile weeks after discharge. A documented, repeatable selection method in place of an uneven one. A governance file that makes the system's next clinical tool easier to do right.

Of the harms listed, which stakeholders are affected?

Patients bear the selection harms, with patients from under-resourced neighborhoods bearing the most under the proxy failure mode, and displaced patients bearing harm invisibly. The transitional care team bears workflow change and list trust. Clinicians bear the tool's effect on their judgment. The system bears the regulatory, financial, and mission harms.

What is the worst plausible case that could emerge from the AI system's development?

The team builds on the extract as it stands, repeat-patient structure and race gaps unexamined, and produces a model whose performance estimates flatter it and whose fairness cannot be certified. The defect is found late, by the review that should have been early.

What is the best plausible case that could emerge from the AI system's development?

Development becomes the occasion to fix the data honestly: the race recording gaps get a documented remediation path, the patient-level structure is handled properly, and evaluation starts from evidence the program can stand behind.

What is the worst plausible case that could emerge from the AI system's deployment?

The tool goes live, selection quietly shifts away from the mission neighborhoods, and the pattern is discovered by an external reviewer or a journalist before the program's own monitoring notices. The system that promised to serve its neighborhoods becomes the case study taught to the next cohort of governance analysts.

What is the best plausible case that could emerge from the AI system's deployment?

The same fifteen percent of capacity, pointed at the right patients: readmissions bending down, override review teaching the model's blind spots, fairness measured and publishable, and a transitional care team that trusts its own worklist. Mrs. Alvarez gets the phone call this time.

SECTION 8: GO / NO-GO DECISION & DOCUMENTED ASSUMPTIONS

Objective: Make and record the gate decision: proceed, proceed with guardrails, or stop. Document the guardrails that condition a "go" and, critically, the key assumptions that would change the answer if the assumptions turn out to be false. This creates an auditable decision trail and gives later stages a set of assumptions to monitor and revisit.

Should we build this?

Proceed, with guardrails. The need is real, the incumbent method is demonstrably inadequate, and the risks identified here are manageable if the sequence is respected: honest data work first, fairness measurement before reliance, human judgment retained throughout. This is a go for the next lifecycle stage, documented data work and review. It is not a deployment decision.

With what guardrails should the AI system be developed?

Document the extract fully before further modeling: the three-way outcome recording and the chosen framing, the repeat-patient structure and how evaluation will respect it, the race recording gaps and their remediation path, the near-empty weight field, and the age banding. Complete the fairness measurement plan, including how utilization features will be tested against the proxy failure mode. Obtain independent review before any reliance.

With what guardrails should the AI system be deployed?

Ranked worklist posture only, per Section 2, with the transitional care team deciding every enrollment. Override logged and reviewed monthly, without penalty to clinical judgment. The displaced-patient picture tracked by group and neighborhood from day one. Nurse training before go-live. A rollback plan that returns selection to current practice.

What assumptions, if false, would change the organization's decision to develop and deploy this AI system?

That the readmission labels are accurate and consistently recorded across ten years. That the extract can be made adequately representative of today's patients. That race recording can be improved enough to measure fairness honestly. That the capacity figures and workflow assumptions from the transitional care team hold. If any of these fails, this assessment returns to the gate.