Impact-Utility-Trust Checklist

For the people inside a health system who decide whether to buy, pilot or scale an AI navigation tool.



A RESOURCE FROM

Scaling Trusted, High-Impact AI Care Navigation

Explore DiMe's full AI Care Navigation Framework →

BEFORE YOU BEGIN

→ **How to think about readiness.** The categories below are a guide for thinking through an AI opportunity. Review them first to gauge what *ready now*, *ready with safeguards*, and *not yet feasible* could look like. Then use the checklist to examine the opportunity more closely. Pay particular attention to the non-negotiables and whether there is a viable funding pathway. The aim is to identify what is in place, what safeguards may be needed, and what would need to change before moving forward.

AI OPPORTUNITY STATUS	*CRITERIA*
Ready now	✔ All non-negotiables met. ✔ Documented friction and a baseline. ✔ A published evaluation of this AI capability at the proposed autonomy level or a more supervised one. ✔ Integration path confirmed. ✔ Monitoring plan with a named owner. ✔ Funding pathway identified.
Ready with safeguards	✔ All non-negotiables met. ✔ The evidence covers only the non-AI version or is early, or the proposed autonomy level goes beyond what the evidence supports.
Not yet feasible	✔ Any non-negotiable unmet; ✔ No funding pathway; or ✔ No evidence combined with high clinical risk.

→ **Does the problem matter to your organization?** Use the Problems that Matter exercise in Section 1 of DiMe's *Playbook: Implementing AI in Healthcare* to test your assumptions. It helps teams define the problem in measurable terms and decide whether AI is the right tool at all.

→ **Is there a plausible path to funding it?** A use case with no funding pathway isn't ready for evaluation. See the Appendix for more details on some of the common funding pathways for care navigation.

NON-NEGOTIABLES

Do not proceed if any of the following apply:

- ✗ There is no viable pathway to HIPAA-compliant integration within the organization's existing or planned infrastructure.
- ✗ A structured clinical safety review has not been completed, or it has identified unmitigated patient safety risks.
- ✗ The capability has no mechanism to preserve human oversight at the points in the care journey where clinical decisions are made.
- ✗ The organization has no capacity, or plan, to monitor this capability's performance post-deployment.
- ✗ Patients have no clear way to reach a person when the AI cannot help them, in their preferred language.
- ✗ The capability's regulatory status (see: FDA CSD (Clinical Decision Support Software)) and applicable AI disclosure requirements have not been determined.

Impact *Is it worth doing?*

If this capability performed as expected, would the outcomes it improves matter enough to patients, clinicians, and the organization to justify the investment and risk? And how strong is the evidence that it will?

Identify which of the following outcome domains this use case is expected to impact. Each is accompanied by an example of established measure(s). Also consider: what is your current baseline for each selected outcome? Over what time horizon will change be visible, and does that match your budget or contract cycle? What is the strongest evidence that this AI capability, not just the non-AI version of the intervention, improves it?

- ☐ **Avoidable utilization:** Fewer emergency visits and avoidable readmissions.
  - 30-day all-cause readmissions (NCQA)
  - Emergency visits per 1,000 patients (NCQA)
- ☐ **Care timeliness and gap closure:** Shorter time to appropriate care, referral reception, and more consistent follow-through on high-risk transitions
  - CMS50, Closing the Referral Loop (eCQI).
- ☐ **Patient experience and navigation ease:** Less confusion, fragmentation, and navigation burden experienced by patients and caregivers
  - The HCAHPS includes care coordination items asking whether staff were up to date and worked well together, and whether the patient and caregiver were prepared for what to do and watch for after discharge (HCAHPS)
  - CAHPS patient experience items on access (AHRQ)
- ☐ **Patient-defined outcomes:** Improved outcomes that patients and families prioritize
  - PROMIS measures of physical, mental and social health, available in more than 80 languages
  - Condition-specific measures like the COST measure of financial toxicity and the NCCN Distress Thermometer
- ☐ **Equity and population-level impact:** Reduced outcome and access disparities, without risk of amplifying them
- ☐ **Lifespan and longitudinal outcomes:** Value that shows up within one fiscal year or accumulates over months, years, or a patient's lifetime.
  - Pair a fast measure with a slow one. At Geisinger, AI-guided colorectal screening outreach raised colonoscopy completion within three months; lower two-year mortality followed, though that finding is a quasi-experimental trial (Park et al., M&SOM 2026).

- [ ] **Care team workload and/or wellbeing:** Less manual chasing, documentation time, or burnout for navigators or clinicians
- [ ] **Safety:** fewer missed escalations, under-triage events, errors in patient-facing information
    - American College of Surgeons Committee on Trauma (ACS-COT) under-triage standard is an adjacent example, not specific to navigation.



Before you move to Utility, you should be able to name the outcome you expect to change, your current baseline, the time frame for change, and the strongest evidence that this AI capability (not just the non-AI version) moves it. If you can't, the capability isn't ready for Utility evaluation.

Utility

Will it actually work?



This is your workflow test. Does the capability address real, documented friction in care navigation for patients moving through your system, and for the teams coordinating their care? A capability can have high potential impact but low utility if it introduces as much friction as it removes.

No capability performs equally across all system types. The right solution addresses a problem that is documented and specific to your patient population and workflows. Use the Unified Patient Journey as a reference set of common points where navigation breaks down.

Use these questions to guide your Utility evaluation:

- [ ] **Navigation friction:** what documented breakdown (a delay, drop-off, redundant step or lost handoff) does this address, and is there current-state evidence that it's real and recurring?
    - Common sources: referral status reports, call center abandonment rates, no-show rates, portal message volume and response times, days from discharge to follow-up, and what patients and caregivers tell you.
    - Published benchmarks can help you judge how big your gap is. In one large academic primary care network, only 34.8% of 103,737 referral scheduling attempts ended in a documented completed appointment. A Penn Medicine trial tested automated texting reminders, but the texting did not raise the rate of patient completing primary care visits within 14 days (32%).
- [ ] **Workflow integration:** does the solution work inside the systems and workflows staff already use, or does it add another platform to manage?

- [ ] **Capacity to act:** is a team in place, with time, to act on what the AI surfaces? A worklist, alert or escalation with no one to respond adds risk without benefit.

- [ ] **Data readiness:** are the data the tool depends on clean and available, such as accurate provider directories, scheduling the tool can write to, and a timely admission, discharge and transfer feed?
    - The Kahn Harmonized Data Quality Framework is widely cited. It has three categories: conformance (does the data match the expected format), completeness (is the data present) and plausibility (do the values make sense).

- [ ] **Patient reach:** Can patients use it in their language and at their literacy level, with accessibility support, and through the channel they prefer, including voice and text, not app only? Do the same functions work in every channel?
    - Use the AHRQ health literacy toolkit as established guidance.

- [ ] **Usability:** Can your intended users, patients, caregivers, and staff, actually use it? DiMe's V3+ framework sets out how to test this: define who the users are and where they'll use the tool, analyze where use errors could cause harm, then test with representative users. V3+ was written for sensor-based digital health technologies, but many of those steps carry over to navigation tools.

- [ ] **Population validity:** Does it work for your patients? Have the models been validated on data that represent your patient population, with results reported by subgroup? DiMe's Playbook recommends asking about fit to your setting and subgroup validation at every stage, from sourcing to contracting.

  "Our work with predictive analytics thus far hasn't been very encouraging. We're finding that the models from our EHR vendor don't perform very well in the communities that we provide care to."

  - Eric Maurer, Chief Innovation and Strategy Officer, Community University Health Care Center (CUHCC)



Before you move to Trust, you should be able to specify which navigation problems this capability addresses, in which workflows, through which infrastructure, and by what mechanism to achieve your desired impact. If those answers are not yet clear, the capability is not ready for Trust evaluation.

Trust *Can we do it responsibly?*

Trust is the foundation of successful deployment, monitoring, and governance and must be maintained throughout a tool's lifecycle. Under what conditions can this AI capability earn the right to operate in the care journey, and what governance structures must be in place before it does? Clinicians must trust that a tool is safe and reliable enough to deploy, and patients must trust that it is serving their preferences and best interests. **Both groups must trust that meaningful human oversight remains in place where necessary.**

Use the following questions to guide your Trust evaluation:

- ☐ **Transparency and explainability:** Can clinicians, coordinators, and/or patients understand what the AI is doing and why?
    - **What good looks like:** Patients can understand what the tool can and cannot do, while clinicians have direct access to the explainability of AI actions or decisions.

      "I have a case study where I used an ambient scribe and it hallucinated that I had prescribed a mood stabilizer for an indication it wasn't indicated for, for PTSD. There was nowhere in the transcript, I was able to go back and check it, that stated that. But ... who at that point is liable if there's a bad outcome? It would fall on me as the human expert in the loop."

      – Dr. Jennifer Shannon, Chief Medical Officer & Co-founder, Glacis; Practicing Child & Adolescent Psychiatrist

- ☐ **Governance structures:** Is there defined monitoring and accountability for AI decisions and failure response?
    - **What good looks like:** A named governance body oversees decision making, safety and performance with accountable owners and clear escalations.
    - Use tools like DiMe's Playbook and governance resources, or the Joint Commission-CHAI guidance on Responsible Use of AI in Healthcare.

- ☐ **Agentic boundaries:** Are clear, enforced limits in place defining what AI can and cannot do autonomously?
    - **What good looks like:** Festor et al. define 5 increasing levels of autonomy of AI decision making applications in healthcare, from (0) no AI involvement to (4) AI making decisions without human intervention.

- ☐ **Equity and bias monitoring:** Does the system actively guard against unequal harm, and is that monitored continuously?
    - **What good looks like:** DiMe's Playbook calls for live performance metrics stratified by race, gender, age and language, with a named model steward and an escalation path to pause, recalibrate or retire the tool. This matches FDA's good machine learning practice principles, which call for representative data and for watching deployed models for unintended bias, and CMSS's call to monitor algorithms for bias.

- [ ] **Patient awareness and informed use:** Are patients appropriately informed when AI is involved, and do they retain the ability to make informed decisions and act accordingly?
    - **What good looks like:** Patients know when AI is involved, can reach a human or opt out without losing access, and are able to correct their information stored by a tool.
    - California's AB 3030 specifies GenAI notification requirements when using GenAI to communicate patient clinical information and CHAI provides broader best practices. DiMe's Playbook also presents what this looks like in practice.

- [ ] **Integration, data, and reliability:** Will the capability perform consistently in your system, on your data, and with the patients you serve?
    - **What good looks like:** DiMe's Playbook covers data mapping, pipeline monitoring and pre-launch validation methods.

- [ ] **Vendor data use**: Is patient data used to train or improve the vendor's models, and is that governed by your agreement?
    - **What good looks like:** Explicit contractual terms and Business Associate Agreements (BAAs) prohibit vendors from retaining, using, or training proprietary models on patient data without explicit opt-in authorization and de-identification, adhering to HHS HIPAA Privacy & Security Rules and the NIST AI Risk Management Framework.

Appendix. Common funding pathways for AI navigation

Coverage varies, confirm locally.

Pathway	Key requirements
Principal Illness Navigation (PIN), G0023/G0024; PIN–Peer Support, G0140/G0146	A serious, high-risk condition expected to last at least 3 months, such as cancer, heart failure, COPD or dementia. An initiating visit by the billing practitioner (not an emergency or inpatient visit). Patient consent before or when services start, and every year after. Trained auxiliary staff working under the billing practitioner (CMS FAQ).
Community Health Integration (CHI), G0019/G0022	Unmet "upstream drivers," such as transportation, housing or food, that significantly limit the practitioner's ability to diagnose or treat the problem addressed in an initiating visit. Documented consent in advance. Only one practitioner per month can bill (CMS FAQ). CMS replaced "social determinants of health" with "upstream drivers" in 2026 (ASH summary of the 2026 fee schedule).
Transitional Care Management (TCM), 99495/99496	Contact with the patient within 2 business days of discharge, by clinical staff who can address the patient's needs, and a face-to-face visit within 7 or 14 days (CMS MLN booklet).
Chronic Care Management (CCM)	Two or more chronic conditions expected to last at least 12 months. An initiating visit for new patients. Documented consent. At least 20 minutes of care coordination a month (CMS MLN booklet). Advanced Primary Care Management codes (G0556–G0558) offer a bundled primary care option (CMS FQHC booklet).
Remote physiologic and therapeutic monitoring	For remote physiologic monitoring: an FDA-defined device, an established patient and patient consent. The 2026 CMS fee schedule added codes for shorter monitoring periods and less management time.
Value-based contracts and quality incentives	Shared savings and quality bonuses tied to measures such as screening completion, readmissions and follow-up after discharge. The business case depends on whether the organization holds the financial risk
Operating budget	No dedicated code. The case rests on staff capacity, filled appointment slots, patient safety and risk reduction. Check vendor savings claims independently.