BRIDGE
Blueprint for Resilient Integration and Deployment of Guided Excellence

Blueprint for Resilient Integration and Deployment of Guided Excellence

Image created by Gemini Advanced 2.0 Flash

With the exception of the images and summaries, this document was not created with the aid of any LLM product for prose or description. It is novel and was created completely by its authors.

“Computers are tools, not rivals...As computers become more and more powerful, they won’t become substitutes to humans. They will become complements.”

Peter Thiel, *Zero to One*, chapter 12

Image generated via Grok3

Table of Contents

Authors....

Joshua Streit,
AVP of Digital Transformation, Aidoc

Dr. Reut Yalon, PhD,
Chief Product Officer, Aidoc

Daniel Spector,
AI Architect, Aidoc

Dr. Jerome Avondo, PhD,
VP of Clinical and Reimbursement, Aidoc

Brad Genereaux,
Global Lead Healthcare Alliances, NVIDIA

Dr. Nina Kottler,
Chief Medical Officer, Radiology Partners

Dr. Stephen Aylward, PhD,
Global Developer Relations, Medical Technologies, NVIDIA

Dr. Alex Towbin,
Associate Chief Medical Information Officer; Associate Chief of Staff; Associate Chief, Department of Radiology (Clinical Operations/Radiology Informatics) Cincinnati Children's

Dr. Efstathia Andrikopoulou,
Echocardiography Medical Director, Harborview Medical Center and Associate Professor of Medicine and Collaborative Intelligence at the University of Washington

Dr. Michael A. Bruno,
Professor of Radiology & Medicine, Vice Chair for Quality and Chief of Emergency Radiology, Penn State University

Yuval Segev,
Chief Information Security Officer, Aidoc

Demetri Giannikopoulos,
Chief Growth Officer, Rad AI

Brenton Hill,
Head of Operations and General Counsel, Coalition for Health AI (CHAI)

Dr. Amit Gupta,
Division Chief Cardiothoracic Imaging, Associate Professor, University Hospitals Cleveland Medical Center/Case Western Reserve University

Dr. Edward Zaragoza,
Vice Chair Radiology Information Technology at UCLA Health, Clinical Professor UCLA Radiology

Ashley Weber,
Vice President of IS Ancillary Services, Ochsner Health

Dr. Leonardo Bittencourt,
Abdominal Radiologist, Vice-Chair of Innovation, Associate Professor, Department of Radiology, University Hospitals and Case Western Reserve University

BRIDGE as a Podcast

If you would prefer to listen to a summary of the contents of this document, just click the above link to get up to speed on BRIDGE in the form of a conversational podcast produced by Google's NotebookLM. Enjoy!

notebookLM

If you'd prefer to just ask your existing questions of BRIDGE, click the above link, upload this document to NotebookLM, and leverage the interactive chat interface. NotebookLM will use the contents of this document as the source material to answer your questions. We hope you find this function helpful in getting directly to your specific points of interest.

TL;DR

Healthcare has been a source of recent and rapid development of AI-driven applications. Just attend any current healthcare conference and peruse the exhibit hall and try to count the number of those present which do NOT display their utilization of some form of artificial intelligence directly on their booth's display. Many of those (+1,000 so far) have achieved FDA/CE Marked clearance for their use. However, the clinical domain, with the exception of concentrated successes in medical imaging, has yet to achieve broad, sweeping standardized adoption. Our contention is, while there are many reasons for this, those obstacles hindering adoption of tools developed thus far can be overcome if they are well understood with broad consensus.

BRIDGE's intention is to outline those attributes which, once appropriate accommodations have been made, should help accelerate both the development and adoption of clinical AI-driven solutions to the betterment of us all.

To that end, the following are a set of statements intended to quickly outline some of the elements related to the more detailed sections below, which can help convey just how high the development bar is for a clinical algorithm to do anything more than become the subject of an academic publication. It is noteworthy that there can be tremendous value to

such publications and some have actually gone on to become entire companies (i.e. Automatic CT perfusion maps by Rapid AI and FFR by Heartflow, both from Stanford). This paper and these statements are focused on how the exception might go about becoming the rule. In order to achieve this level of "graduation" from idea to publication to product to production use, the bar is achievable but decidedly high. To allow development of this kind to reach broader application, one should understand the following and make the appropriate accommodation:

The costs to implement a single, clinical AI solution in a production environment and support it over time can exceed **$200,000 per solution,** not including any regulatory approval related costs(1).

Development cycles to create a solution often occur over a 12-24 month period of time (or more).

The dataset necessary for a given clinical use case is often at least 100,000 instances of that particular use case.

The dataset must be at least as diverse - in patients, modalities, manufacturers, protocols, and varying study quality used to produce the data - as the population to which a given solution is proposed to apply.

If it is an imaging solution, the imaging exams in question have to be both high and low quality in order to function well in a production environment which will be made of a similar range of exam quality produced in real time. It must also be produced from a plurality of modality vendors as the images they create have subtle (sometimes not subtle) variations.

The data collection budget to produce a single imaging solution can exceed $1 million.

The computational budget to produce a single imaging solution can exceed $1 million.

The engineering budget to produce a single imaging solution can exceed $1 million.

Any AI-solution will have a set of minimum viable production environment characteristics which must be available or the algorithm-driven solution will fail to produce the desired outcome or output. These attributes are related to but still very different from the algorithm itself.

The regulatory requirements are significant but known and also can cost between $50,000 - $1 million.

Achieving regulatory approval from Federal agencies often takes +12 months, starting from after the curation and internal testing of a given AI solution.

The difference between a model and a solution is this: one is a math equation (or set of equations) designed to present precise data, the other can change the course of many peoples' lives (clinicians and patients alike).

There is no current broad adoption for reimbursement for the use of any AI tool. There is lots of activity, however, in an effort to make this come to fruition.

Change management will continue to be critical for the success of the wide variety of transitions needed to adopt AI driven solutions. Without a thoughtful approach here, even well crafted models can suffer from poor adoption and utilization.

If the workflow which triggers an AI solution isn't predictable, reliable, measurable, and completely and totally automated, its chances of failure are high.

If the solution being proposed requires EMR data and/or needs to be used within the EMR and you don't already have a well understood, mature FHIR compatibility, stop. Fix that. Then start over. Creating an accurate AI algorithm is only part of the solution.

If the proposed solution is one involving medical imaging, one must understand the implications of the creation, visualization, distribution, and long term storage of AI generated outputs as a core function of the overall solution.

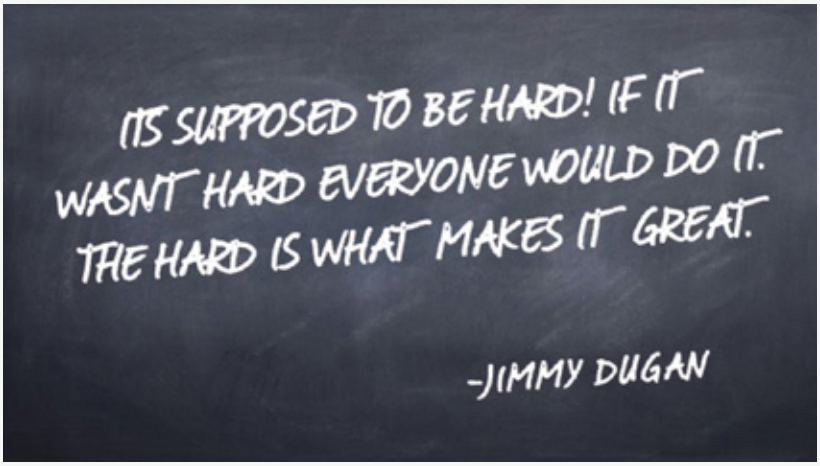
The only path to financial return for clinical AI solutions is better patient care.

But no one pays for quality alone.

Summary:
Creating or buying clinical AI solutions can be a lengthy and expensive endeavor. However, executing well can yield significant and dramatic benefits for clinicians and patients alike. One would do well, therefore, to make certain the following are in place as one begins a journey of either acquisition or creation:

- Depending on the scale of one's organization, one should expect the necessary budget to be seven figures.
- Plan a 24 month success cycle from ideation to implementation and assessment. Engage with one's legal and regulatory team early in the ideation process so the project is clear on its regulatory path.
- Ensure the data set to be used for training is large and diverse in both how it was created and the patients from whom it was created.
- Have a well outlined and understood method by which one's targeted users will consume AI generated data in a manner inline with existing workflows, user interfaces, and regulatory requirements.
- Nothing implemented can rely on a manual process of any kind.
- All technical methods needed to access, consume, analyze, and output the data type in question need to be thoroughly understood.

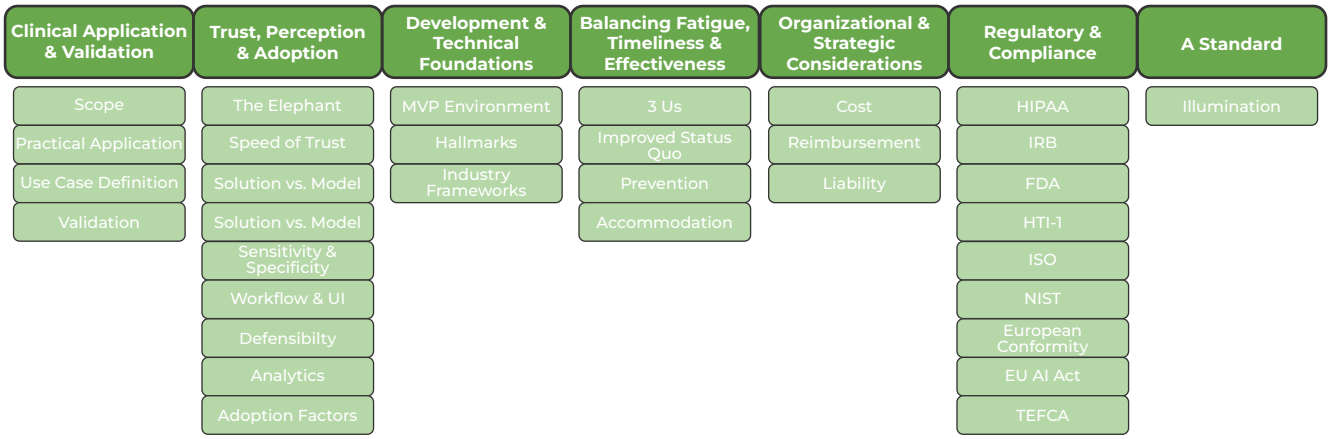
We know. That all sounds pretty difficult. It is. But therein lies its value. To borrow a line from a Tom Hanks movie, A League of Their Own: "The hard is what makes it great." Whatever one endeavors to create to truly advance another's clinical practice, will no doubt require significant resources, time, and effort. It is our belief that together we can accomplish this lofty outcome. With the global community of like minded innovators, we can build a BRIDGE into that future.



1. Clinical Application and Validation

1.1. Intended scope of the BRIDGE Framework

Photo of the Brooklyn Bridge by Joshua Streit in April 2023. Designed by James A. Roebling in 1867.



In the initial release of BRIDGE, we are focused on clinical workflow artificial intelligence (AI) applications. “Clinical Workflow” is defined in these circumstances as the workflow associated with a clinician’s clinical production environment. Therefore these AI applications (or solutions) are assumed to be created, deployed, integrated, monitored, and tuned with this workflow environment in mind.

Existing use-cases that will be frequently referenced include solutions that work with these data types:

- Codified (Structured) and unstructured textual data from a variety of data sources including those in the non-exhaustive list below:

- Electronic Medical/Health Record (EMR, EHR)
- Laboratory Information System (LIS)
- Anatomical Pathology LIS (APLIS)
- Radiology Information System (RIS)
- Cardiovascular Information System (CVIS)
- Genomic databases
- Non-imaging data from Picture Archiving and Communication System (PACS) or vendor neutral archives (VNA)
- Dictation reporting systems

- Image based data:

  - Enterprise PACS and VNA
    - Enterprise Imaging
    - Radiology
    - Cardiology
    - Pathology
    - Ophthalmology
    - Perioperative Video
    - Mobile captured photos and videos
    - Medical photography

  - Imaging modalities or devices
    - CT
    - MRI
    - X-Ray
    - Ultrasound including point-of-care and B-scan devices
    - Nuclear medicine and PET imaging
    - Fluoroscopy and angiography units
    - Whole slide scanners
    - Corneal Topography
    - Visual Field
    - Slit Lamp
    - Optos Fundus
    - Optical Coherence Tomography
    - Perioperative scopes

- Machine Generated Data (MGD): Data automatically collected by medical devices, sensors, and other systems without direct human input. This data can be acquired ad-hoc in the setting of a visit or continuously via a take-home device or in a hospital based setting with connected devices.
    - Waveform: Standard communications protocol for computer assisted electrocardiography (SCP-ECG), PNG, or DICOM
    - Biomarker data: Heart rate, blood pressure, O2, etc…
    - Of note ASTP/ONC previously **provided guidance** for "Patient Generated Health Data" (PGHD) with a focus on wearables

- Relevant Claims Data (where accessible)

- Utilization, engagement, and adjudication data

The forms of AI built on these data types are manifold and could leverage individual or combinations of generative or non-generative AI Analysis methods. Importantly the focus of this framework assumes integration of clinical team members with the relevant above information for ultimate decision making.

While potentially having relevance (such as data bias considerations) administrative and operational AI use-cases are not included as an area of focus for this framework. BRIDGE is solely focused on the clinical care delivery domain.

1.2. Practical Application and Intended Use

The value of an AI solution will be directly related to how your users perceive its impact to be. Additionally, this AI-driven output must be perceived to be greater than what it took to create it in the first place. This perception - imagined or measured - will determine the adoption of what has been deployed. Therefore, strong consideration should be made to help ensure the size of the impact to the end user and, by proxy, the lives of the patients for whom they care is more valuable than the input it takes to create, implement, integrate, train, and measure the tool. The application is often chosen by the leadership of its users and not the end user. Therefore, care should be taken to ensure that the AI model performs as intended from its inception. From this perspective, with each novel AI tool deployed, the effort for the end user to leverage that tool must be less than the outcome gained from its use.

For example, if a solution is designed to identify a pathology on an image, it must do so with a specific outcome in mind (i.e., enhance the detection of that finding vs. that in the absence of that solution) so that a specific goal or goals may be achieved. Additionally, as the output is measured and assessed, it will be constantly assessed against whatever the effort was for a user to interact with it. The balance of this input/output equation from the perspective of the end user will be of paramount focus from their point of view as a leading indicator of the success or failure of each AI-driven implementation.

1.3 Use Case Definition

A given solution will be valued according to its ability to balance many things, hence this framework. Use Case Definition, in addition to functioning in parallel to other use cases and fitting into an appropriate regulatory framework, will be subject to several other aspects which, thankfully, are more pragmatic and straightforward than some of the topics addressed thus far. Below are a few examples of these attributes and an explanation of each. This not intended to be an exhaustive list, but rather to provide one with a solid grounding on the common attributes which comprise an AI-driven use case.

1.3.1. Acuity of setting

The optimal time interval for the availability of the output of a use-case as compared to the timing of the data availability is one of the most important aspects of an AI driven solution. The acuity of setting in which the patient presents is one of the most significant contributing factors in the delivery of AI related output.

The most extreme version of this active in today's market are the existing AI products designed to address the needs involving an ischemic stroke. The timeliness for any treatment for this particular neurovascular disease is measured to the second. Quite literally, if a solution related to ischemic stroke care took an extra couple of minutes for initiation of process in its workflow design, the impact to the patient can be millions of unrecoverable neurons and loss of function or life. Historically there has been a bit of confirmation bias that stroke patients are always in the emergency setting, where some AI solutions have already demonstrated meaningful value. However, there are rare but still substantial patient presentations of patients with hemorrhagic or ischemic stroke in other environments like outpatient care or inpatient settings. AI acceleration offers an opportunity for these atypical presentations in a non-emergent care setting to be treated as emergent cases and reduce variability of care based on setting.

Another classic example of an acute pathology that presents in a variety of settings and could benefit from AI acceleration are clinical degradation or Sepsis protocols. Many relevant pieces of information are flowing in and once an AI solution has enough signal with even limited information escalation can offer significant patient outcomes.

By comparison, if one is designing something to help with the laborious task of constantly synthesizing voluminous amounts of patient history so that its synopsis can be available at the ready for a primary care physician when seeing patients in his/her clinic, the acuity is usually dramatically less severe and, therefore, forgiving when it comes to the timeliness of its feedback compared to the availability of the underlying patient data. As long as the results output is available prior to the scheduled meeting.

However, if the patient history synthesizing function being designed is targeted at a stroke patient, then one is forced to optimize that clinical AI delivery to the point of the most acute care scenario one may ever face - even though it is, essentially, the same function being delivered to a primary care physician in the prior example.

Most importantly the analysis timeliness parameters of the algorithm must be designed based on the underlying acuity of the targeted intervention and not tied to the acuity of the setting. Therefore, a precise understanding of the clinical impact and severity of the patient

acuity pertaining to the AI solution in mind must be clearly understood. As is pointed out above, a highly useful application can be rendered useless if it is implemented into a workflow and/or pathology demand which isn't flexible enough to accommodate the solution as it is designed.

1.3.2. Intended Users

There are multiple factors that influence the selection of an intended user for an AI solution. From a business standpoint, expanding the base of potential users can enhance marketability and adoption. However, from a clinical and regulatory perspective, the intended user should be the individual best equipped to interpret the AI outputs responsibly and accurately. Ideally, this user should demonstrate minimal susceptibility to common human-computer interaction biases, as well as any latent biases embedded in the AI model's training data. In practice, the optimal end-user is someone who can appropriately disregard false positives and false negatives, while confidently acting on true findings.

From the field of medical imaging, emerging draft guidance from the American College of Radiology (ACR) defines a "qualified end-user" as a physician who can independently perform the diagnostic task supported by the AI and who has been trained to assess the validity of the AI's outputs (2).

Other important considerations include the availability of the intended end-user and the clinical context in which the AI output is delivered. For example, in acute scenarios, such as trauma or emergent vascular events, physicians or advanced practice providers (APPs) are often the primary users due to the urgency of decision-making. In less acute or longitudinal care pathways, such as screening or surveillance programs, other clinical roles, including RNs, navigators, or care coordinators, may engage more frequently with the AI outputs, as long as oversight mechanisms are in place to involve a qualified end-user when necessary.

For instance, a good example of how one must factor in their intended users into a full solution is with respect to an aneurysm. The initial presentation of this pathology will likely be to a radiologist. However, if it is above a certain combination of size and patient comorbidities, then a surgeon may need to receive that information to give immediate consultation. Further, if it is below a certain threshold in size and lacks comorbidities, other roles like a navigator will be engaged. Depending on the scope of practice these solutions may be staffed by an RN or APP so the patient can be engaged in a longitudinal screening and surveillance program. The goal of these programs is often to identify patients who might be at a higher risk of an adverse event. Therefore, with just one pathology, the intended users can range from surgeons, radiologists, critical care physicians, RNs, APPs, and entire

departmental support staffs to help maintain patient engagement while its acuity can range from hyper acute to not acute at all.

Ultimately, the breadth of intended users of a given solution underscores the importance of providing a range of local integrations. We will explore this concept in more detail in Section 3.1.6.. The key point, however, is that delivery of AI results to this diverse set of potential users must align to where the end users are doing the majority of their work.

1.3.3. Output Format and Distribution

Formatting is a topic on which a great deal of creativity can be applied. However, within a production, healthcare environment, one can only leverage the formats compatible with the existing clinical systems or endpoints available if one's solution is going to be adopted to achieve scale. So while an AI model may help to generate charts, diagrams, annotated images, secondary captured images, pdfs, reports of all kinds, contributions to other reports...none of all that rich information can be consumed if it cannot be distributed appropriately to its target audience in a way that is easy for them to consume. This further limitation on distribution method(s) mechanically restricts how and to whom this data can flow. Here are a few examples of distribution methods and their compatible data formats which could be leveraged for an AI use case:

- **HL7v2** - .The most common method of exchanging clinical data between healthcare IT systems, which often requires a significant amount of custom mapping on a per system basis. Typical formats for AI results include:
    - Observation (ORU message)
    - Order (ORM message)
    - Document (MDM message)
- **FHIR** - a REST API that follows a standard data schema which is typically used by EMR systems , enabling reading and creating clinical records by external systems.
- **DICOM** - One of the few universally accepted healthcare standards. Focused on medical imaging in Radiology and Cardiology but quickly growing into other domains like Ophthalmology, Pathology, and even waveform data from device generated outputs.
- **DICOM SR** - a subset of DICOM offering structured report associated with DICOM images - Heavily adopted in Cardiology and OB Ultrasounds but gaining steam in radiology.
- Proprietary REST API with specific systems (eg. reporting systems)

Sometimes the inability to integrate or for downstream systems to appropriately display requires a novel user interface but this should be avoided whenever possible.

1.3.4. Classification of Data Output

There are two main points of focus with respect to the data which is derived from an AI solution. That is, does this model output appropriately contribute to the patients' record in some manner or does it not? If it does, then the data will fall into a set of workflow options appropriate for the contribution to the patients' record. If it does not, then it should follow an alternate path of workflow options. Let's look at some examples...

If one chooses to produce data intended for the formation of the actual diagnosis of a patient's condition/presentation, then that solution should likely be considered as an appropriate component of the patient's record. It, therefore, is mostly likely to be reviewed within that record, which is typically one of the systems covered in section 1.3.5., depending on the data. This places further constraints on the workflow which is possible with this data as the clinical environment must be able to support the needs of that use case.

> ***NOTE***: *if a solution is contributing directly to the diagnosis of the patient, please make the appropriate considerations with respect to the regulatory and governance framework which is applied to the solution. A regulatory overview is covered in Section 6.*

Conversely, if the output of a solution or model is not intended to contribute directly to a patient's record, then it should very likely be consumed/interacted with, etc. via some alternative application in addition to that which controls his/her record. This is where novel workflow solutions can be considered like:

- Mobile interface
- Overlay/sidecar applications
- Web Interface

From these supportive interfaces, strong emphasis should be placed on driving the efficiency of the end users involved. Consideration should be given to the fact that they already spend the vast majority of their time embedded within the clinical record systems. Therefore, introducing new data via an alternative interface, will need to be delivered in such a manner that augments their workflow such that it is worth the disruption from the clinical applications in which they are already immersed.

A straightforward example of this in use in production already is within medical imaging in which there are already a variety of FDA approval levels for those algorithms which can be mapped to the inclusion/exclusion within the patient's medical record. Here are some examples of those approval levels and the recommended inclusion within the medical record:

- **CADt** - triage algorithm - the output is the “triage results” which is a yes, no, not analyzed, or error status. Any other outputs are considered incidental.
- **CADe** - detection algorithm, Class II - For results reviewed and confirmed by the intended user it should be documented in the chart.
- **CADx** - diagnostic algorithm, Class II or III - For results reviewed and confirmed by the intended user it should be documented in the chart.
- **Measurement algorithms** - Class I or II - For results reviewed and confirmed by the intended user it should be documented in the chart or report. If image based the record of measurement should be retained in the appropriate xPACS. Note this may be a reannotation by the end user.

> ***NOTE***: *One should always consult appropriate local teams for the solution being considered with respect to inclusion/exclusion of the output from the solution into the long term clinical record (EMR/PACS).*

Finally, one should also be mindful that, no matter the regulatory classification or the inclusion/exclusion choice of the solution output, due to the Parallelism of these solutions and/or the fact that even those covered immediately above and more could be used in combinations in order to fulfill a clinical or workflow need. The intended use, output format of the data, appropriate regulatory classification, inclusion/exclusion criteria, how multiple solutions are used in combination with one another etc. are all factors which are going to contribute to where and how your targeted end users will be able to consume or interact with the product of your solution.

1.3.5. Triage and Acceleration

One of the most immediate, positive contributions that AI can have in today’s healthcare workforce is simply to provide a means of more appropriately applying its own scarce resource to the patients they treat each day. This dramatic need is part of why so many of today’s CE Marked or FDA cleared solutions fall into the Triage category of algorithms. One might consider it a lower hanging piece of AI fruit in which one's organization can benefit greatly from being able to leverage its existing resources in a new or more intelligent manner so as to achieve a higher output and/or quality for its patients.

1.3.6. Patient Facing Record

Finally with respect to use case definition and its contributing factors, one should be mindful of whether or not a patient will end up with access to the output from an AI solution with the understanding that it was generated from such a source. One of the more obvious examples of this would be when they have access to their own imaging which has been annotated by a model which has had access to their study. In the case of discrepancies between the written record and AI Analysis record this could result in confusion from these patients and their engaged post-event clinical team.

Additional information, regarding factors related to Use Case definition are provided by Gemini Advanced v1.5 Pro Deep Research and **can be found here**.

1.4. Validation - Looking Outside the Box

> *"Be passionate about solving the problem, not proving your solution."*
>
> **Nathan Furr, 5x author, INSEAD faculty member**

Image created by Gemini Advanced 2.0 Flash, "Looking outside the box for model validation"

1.4.1. The Role of Continuous Model Validation

The validation of clinical AI models is a cornerstone of ensuring ongoing safety, reliability, and clinical utility. As AI applications proliferate, rigorous validation frameworks are essential to assess model performance in real-world clinical environments. Beyond initial accuracy metrics, validation must account for generalizability across diverse patient populations, integration within existing workflows, and longitudinal monitoring to detect performance drift. A major challenge lies in the fact that real-world patients may differ significantly from those included in internal or even external validation datasets due to shifts in demographics, disease prevalence, or comorbidities. Additionally, healthcare delivery evolves over time, with changes in clinical guidelines, diagnostic protocols, therapeutic approaches, and resource allocation influencing model performance. Factors such as data acquisition processes, imaging techniques, or EHR configurations may also shift, necessitating ongoing reassessment. Without a structured, continuous validation approach, AI models risk becoming outdated, misaligned with clinical practice, or even introducing unintended biases that could compromise patient care.

To address these challenges, we propose an iterative validation framework inspired by Quality Improvement (QI) methodologies, particularly the Plan-Do-Study-Act (PDSA) cycle. Unlike static validation paradigms, which assess models at a single time point, this approach ensures continuous model refinement and adaptation based on real-world feedback. Each cycle begins with:

Planning establishing validation metrics and identifying emergining risks	**Implementation** monitoring model performance in practice
Assessment analyzing discrepancies, including false pos/neg and provider interactions	**Refinement** adapting models, retraining on new datasets, or modifying workflows

This cycle should be embedded into hospital AI governance structures to ensure that AI tools remain clinically relevant, safe, and equitable over time. Additional considerations include changes in patient and/or clinician engagement, as adoption of AI tools may fluctuate based on trust, usability, or accessibility, and variability in institutional resources, which may

influence the sustainability of AI-driven solutions. By implementing a dynamic, feedback-driven validation system, healthcare organizations can maximize AI's clinical impact while maintaining trust, transparency, and patient-centered care.

1.4.2. Explainability and Interpretability as Core Components

As AI systems become more autonomous and sophisticated, explainability and interpretability will be fundamental to their validation and adoption in clinical practice. Unlike traditional clinical tools, many AI models—particularly deep learning algorithms—operate as "black boxes," generating outputs without readily interpretable reasoning. This opacity poses a major barrier to trust and adoption, as clinicians require transparency in decision-making to confidently integrate AI-driven insights into patient care. Therefore, explainability must be an explicit component of model validation, ensuring that users can understand how and why an AI system generates a particular recommendation. This extends beyond performance metrics to include clear justifications for predictions, visualization of key decision-driving features, and mechanisms that allow clinicians to probe and challenge AI outputs.

A crucial component of continuous AI validation will be establishing clinician-in-the-loop workflows, where users can interact with AI models, adjust parameters, and provide real-time feedback on model outputs. This will enable healthcare institutions to systematically track instances where AI recommendations conflict with clinical intuition or best practices, triggering additional validation cycles. Techniques such as saliency mapping for imaging AI, attention heatmaps for text-based models, and probability calibration curves can help translate complex AI outputs into human-understandable formats. Additionally, embedding confidence scores and risk stratification explanations into AI-generated reports can enhance usability while mitigating the risk of over-reliance on AI-generated outputs.

Regulatory frameworks such as the FDA's Software as a Medical Device (SaMD) (outlined in the Regulatory and Compliance section) guidelines and the HTI-1 regulations increasingly emphasize transparency, meaning that AI developers will need to incorporate interpretability by design rather than as an afterthought. Furthermore, institutional AI governance structures should mandate explainability audits as part of routine AI performance monitoring, ensuring that clinicians and patients alike can question, refine, and trust AI-driven insights. By prioritizing interpretability as a validation metric, AI solutions can better align with real-world clinical decision-making, fostering a culture of shared decision-making rather than blind acceptance of algorithmic outputs.

1.4.3. In-situ Validation of Performance

Transitioning from a controlled research environment to the dynamic clinical setting necessitates in-situ validation. Imagine a deep learning model that predicts hospital readmission risk. During development, it may be trained on a curated dataset with complete electronic health records. However, in real-world use, it might encounter missing lab values or inconsistent documentation. Hence, depending on one's use case, significant consideration may need to be given to the inclusion of low and high quality data to be included within the training data set in order to be able to address real world scenarios. Additionally, continuous monitoring is essential to ensure the AI maintains its performance as it is then subject to data upon which it was not trained. This could involve comparing the AI's predicted readmission risks with actual outcomes and flagging discrepancies for review by clinicians, which is then a workflow that would need to be designed and deployed along with a model in such a manner that it is convenient for the end user. This approach can be applied to both imaging as well as clinical based use cases with significant consideration to the workflow impact being made with the appropriate monitoring and/or feedback capabilities provided.

1.4.4. Initial Calibration & Proxies for Performance

Before deploying any AI, meticulous calibration is essential. Consider an AI that flags potential cancerous lesions in mammograms. The threshold for flagging a lesion as suspicious needs to be carefully calibrated to balance sensitivity (detecting true cancers) and specificity (avoiding false alarms). This calibration might need to be adjusted based on the specific population being screened and the preferences of the radiologists. However, continuously evaluating every AI decision against a gold standard can be impractical. In this case, a proxy for performance might be the rate at which the AI flags lesions that are subsequently biopsied. A significant drop in the biopsy rate without a corresponding increase in missed cancers could indicate that the AI is effectively helping radiologists prioritize suspicious cases. Delivering this full solution would need one to be able to deploy an imaging workflow which was effective for the interpretive process as well as a multimodal mitigation workflow to be continuously run in the background complete with an adjudication process for discrepant cases which could have clinical impact. These discrepancies could then be leveraged for additional adaptation of the model deployed.

1.4.5. Human + Artificial Intelligence

Combining the benefits of both human and artificial intelligence highlights the synergistic potential of combining the out of both. An AI system that analyzes pathology slides can identify subtle cellular features that might be missed by the human eye, but the pathologist's expertise is crucial for interpreting these findings in the context of the patient's clinical history and other diagnostic information. Commonly in radiology, there are unexpected/incidental findings which are not related to the indication for the order in question. Radiologists are trained to primarily address the clinical question driven for the reason for the study. Effective imaging algorithms can help radiologists catch both the expected and the unexpected pathology for their patients. Similarly, an NLP model can summarize lengthy patient records, but the physician's empathy and communication skills are essential for building rapport with the patient and understanding their individual needs and preferences. In each of these examples, the output goal of these models is to augment existing work which solely rests on the shoulders of today's clinicians.

1.4.6. Ongoing Mitigation

Even the most rigorously validated AI systems can make mistakes. Imagine an AI that recommends insulin dosages for diabetic patients. If the AI encounters a patient with an unusual medical condition not represented in its training data, it might generate an incorrect dosage. Mitigation strategies are crucial. This could involve setting limits on the AI's recommended dosage adjustments, requiring clinician approval for high-risk recommendations, and implementing alerts for unusual patient profiles. Regular audits can identify systematic biases or errors, ensuring the AI remains safe and effective.

1.4.7. When 90% isn't good enough

In clinical settings, even small errors can have significant consequences. An AI that predicts sepsis risk with 90% accuracy might seem impressive, but the 10% of missed cases could represent patients who deteriorate rapidly without timely intervention. Similarly, a medical imaging AI that misdiagnoses a brain bleed even a small percentage of the time could have devastating consequences. We must strive for levels of accuracy that far exceed typical benchmarks, continuously learning from real-world data and feedback to improve the AI's performance. As is mentioned in section 2 (Trust, Perception, and Adoption), establishing and maintaining the trust of one's end users is of paramount concern. A general rule of thumb in this regard to keep in mind is that a solution needs to be as effective as a clinician would otherwise be on their own if given adequate data access and synthesis time. If a solution isn't statistically as good as the clinicians being served, those clinicians are very likely to lose trust and reject the solution in question.

1.4.8. Unexpected Opportunities for Enhancement

Deploying AI in clinical practice often reveals unexpected opportunities for improvement. For instance, an AI system designed to assist with triage in the emergency department might reveal bottlenecks in the existing workflow or highlight the need for better communication between nurses and physicians. Similarly, an AI that analyzes electronic health records might uncover previously unrecognized patterns in disease progression or treatment response, leading to new research hypotheses and clinical trials. Hence, one should fully expect to be presented with adjacent or new challenges to solve once one's primary objective has been adequately addressed. For instance, one may leverage a transformer NLP model to extract existing diagnosis from the clinical record which need to be addressed by a specialist. Once those cases are presented, that specialist may need additional augmentation with respect to segmentation or measurement of any related imaging which was acquired related to the original diagnosis.

In summary, AI validation must move beyond one-time approval processes toward a continuous, iterative improvement cycle modeled after QI practices. This approach acknowledges the dynamic nature of healthcare and recognizes that AI models must evolve alongside shifting patient populations, clinical guidelines, and resource constraints. Furthermore, explainability and clinician engagement must be prioritized, ensuring that AI is not just an advanced computational tool but a transparent, accountable partner in patient care. By embedding these principles into AI governance, healthcare systems can maximize AI’s potential while safeguarding clinical integrity, trust, and equity.

While metrics like accuracy, sensitivity, and specificity are essential, clinical AI demands a more nuanced approach to validation. For example, an AI system for detecting pneumonia in chest X-rays might achieve high overall accuracy but struggle with specific demographics, like pediatric patients or individuals with underlying lung conditions. Similarly, a natural language processing (NLP) model designed to extract medication information from clinical notes might be less accurate for patients with complex medication regimens or those who see multiple providers. One must go beyond aggregate metrics and evaluate performance across diverse subgroups, disease severities, and data quality scenarios. Analyzing "failure modes" is also critical. Perhaps the pneumonia detection AI tends to misinterpret chest tubes as signs of infection, or the NLP model struggles with abbreviations and negations. Understanding these weaknesses allows for targeted improvements and informs clinicians about when to exercise caution.

2. Trust, Perception, and Adoption

2.1 Addressing the Elephant in the Room

Over the last few years, one no doubt has noticed the dramatic increase in the creation and promotion of imaging-based AI solutions. In 2023 alone (the last complete year of data available as of this writing), there **were 222 new FDA** clearances (3). As of September 2024, 776 of the total 1018 medical imaging specific FDA approvals (which date back to 1995) are within radiology (3). A similar phenomenon has been observed outside the US via the number of CE Marked solutions in 2024 has continued a steady climb. There is a noteworthy difference between the US and European approach to keeping track of the total number of solutions receiving regulatory approval. That is, the CE Marked (covered in more detail within the regulatory section of this document) process does not keep track of a central repository of the total number of solutions approved. However, it is believed that the total number of solutions between the FDA and the European Commission to be roughly equivalent.

Image created by Gemini Advanced 2.0 Flash: Elephant in the Room

A confluence of an accessible regulatory framework, dramatic clinical need, large volume of available training data, and functional, iterative technology available at scale (GPUs, cloud computing, etc.) have all led to algorithm proliferation.

When assessing cleared algorithms, a common set of themes emerges:

- Their intent is to help improve the quality of care while reducing the burden or tedium of work for the end user
- There are a wide variety of use cases covered by those approved solutions
- They all need the same kinds of supportive functions to help ensure their success

Driving quality, the chosen use cases, and the wide variety of tasks which need to be addressed for each use case are all aspects BRIDGE addresses throughout the remainder of this document.

Later to the party is Generative AI, which is currently limited to language based use cases, but may not in the very near future given that several Vision Language Models (VLM - similar to LLMs but can understand images, etc.) have entered validation phases. Of these being used in production are transformer based NLP models, some which may not require FDA clearance because they do not meet the FDA's definition of Software as a Medical Device (SaMD) (2). However, any AI model deployed within the EMR, including language models which may not be under the FDA's scope as a medical device, now falls under the Office of the National Coordinator's HTI1 and the EU AI Act (both covered in a later, regulatory section), muddying the waters further.

Due to the limitation of how much data NLP can handle accurately and the comparatively nascent stage of LLM related clinical projects, most have opted to focus on imaging use cases. One should be mindful of these considerations as they will want to balance a variety of needs as they consider what type of project will be in their best interest to pursue.

Today, we have a proliferation of imaging-related solutions which need to be balanced against any of those one may also choose to contribute. As will be covered throughout this document, there are a wide variety of factors which will contribute to the success, or its limitation, of a given project. Not the least of these will be how well the creator of a novel solution can fold in the outcome of their efforts into what could be a kaleidoscope of solutions already in production (or will be in the near term) within one's clinical environment.

For a deep dive regarding the FDA's Approvals and Trends for Artificial Intelligence in Medical Imaging, please see this **Summary Document**.

2.2. Implementing AI at the Speed of Trust

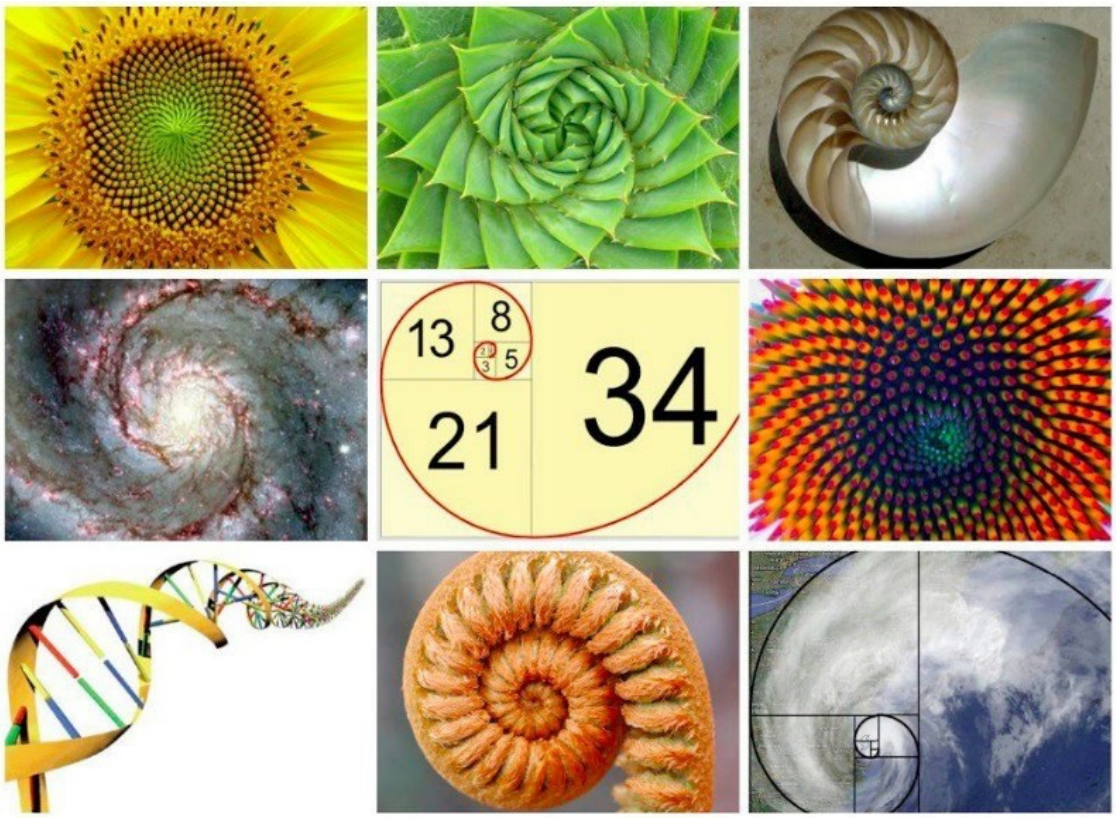
> *"Earn trust, earn trust, earn trust. Then you can worry about the rest."*
>
> **Seth Godin, 20x author, February 23, 2014**

There is no doubt that AI plays a key role in healthcare. It is important to remind ourselves "why" this is so, and what it is and isn't: to separate hype from reality. Healthcare is the ultimate service industry, where life itself is directly impacted by action and inaction. AI is not a replacement or an alternative to this industry, it is a tool which aids those delivering care with insights. There has been plenty of distraction, but as time has worn on, it is clear these technologies have the capability to be used in tandem with traditional approaches to provide an enhanced experience for the patient and the clinician alike (4).

There are, however, a number of challenges which have played a role in slowing and, in some instances, preventing clinical AI adoption into production (used under real world conditions on live patients). Most of those challenges revolve around the concept of trust. Meaning, how can a solution build the required degree of trust with an institution, individual provider, or even a patient to the point of complete adoption? Hesitation in jumping directly into this pool is a natural and appropriate human reaction to any novel approach or technology. It is especially so when that approach or technology applies to healthcare where any decision will potentially impact human lives.

In order to overcome this natural barrier of entry to production/clinical use, one must, therefore, develop and deploy the necessary means of creating such a standard of trust. This standard will have a number of components or layers to how the trust of end users can be achieved, and combinations of the attributes making up these layers will need to be accounted for in order to meet such a high bar in the mind(s) of the targeted end users.

This first section will outline important layers as well as the means by which they can be addressed, thereby beginning the process of building out an infrastructure which can be trusted to deploy an array of AI-driven solutions. In an ideal setting, with each successive use case deployed, the trust should compound, allowing the rate of adoption of those solutions to increase over time in a kind of Fibonacci (Golden Ratio 1.618) accelerating spiral effect we find elsewhere in nature. This can then become a process by which each organization can establish their velocity of trust which continues a path of acceleration as more and more use cases are applied to the same structure. Just as with AI itself, the more data subjected to a model, the better that model is then able to perform, so it can be with your respective team(s). The better they get at structuring the manner by which AI solutions are designed, implemented, and maintained, the more efficient and quickly they will be able to do so in a manner on which they can begin to rely.



Examples of the Golden Ratio in nature.

Notably, strategic selection of the first batch of use-cases will often determine long-term success in this standard of trust-building practice. An initial use-case with the appropriate mixture of prevalence, outcomes, (healthcare organizational and patient-related) cost, and return on investment (for the organization, as well as the clinician and patients) will ultimately play a large role in the organization's perception of this kind of solution.

The first aspects related to AI deployment which materially impact the speed of adoption and trust are the following:

- Mistaking a model for a solution
- Prevalence and perception
- Accuracy and specificity
- Workflow and user interface
- Defensibility of results
- Practical application of intended use
- Analytics related to each use case and its users
- Change Management associated with new capabilities

2.3. Solution vs Model

1. A full "Solution"

2. A collection of "models"

The latter (2) is needed but in no way represents the former (1). Both images were generated by Gemini Advanced 2.0 Flash.

A model generates data whereas a solution generates outcomes. This isn't to say a model by itself does not generate value but, without considerations from solution ideation of how the end user could or should interact with the output generated by the model, healthcare organizations may find themselves in a scenario in which the model merely fits their solutions. At the user interaction level, a successful principle is one of Radically Integrated Transformation. Meaning, the author of a novel solution will need to both exhaust all possibilities of delivering model output within an existing user interface as well as make efficient use of the health systems data which needs analysis by their model. The design and implementation of such 'Radical' solutions are founded on the principle of respecting the native environment and workflow of the end-user as much as possible. This approach allows maximization of the clinical context and decision-making. To execute on this perspective, one will often need access to multiple data points, stored/generated from multiple sources, in order to synthesize the needed information. An algorithm or model, in these circumstances, is a component of rather than a total solution. Developing and deploying the latter requires an ever growing number of native integrations in order to be delivered as a 'solution.'

As a general rule of thumb: integrate aggressively, but also don't be afraid to transform workflow if there is additional value to be gained by augmenting with new interaction

methods. These novel interaction methods should be limited to scenarios where native integration is either not possible or not robust enough to deliver the full value or requirements of the solution.

Modern interoperability protocols like Substitutable Medical Applications Reusable Technologies (SMART) standard on HL7's Fast Healthcare Interoperability Resources (FHIR) (collectively SMART on FHIR) and FHIR more broadly have enabled new levels of interaction with in situ systems. However even these integrations are ultimately limited by the systems in which they find themselves located. The barrier to entry of "yet another app" is also very real as clinical teams find themselves overwhelmed with mobile apps, desktop apps, and web pages. In each solution delivery method consider whether existing modes of delivery (EHR, Information System (xIS)), Visualization system (xPACS), and AI Platforms are able to support the level of interaction necessary for end-user satisfaction and utilization. Another key aspect of integration with existing systems is that new solutions can benefit from the "halo effect" of others solutions delivered through those methods that may have more frequent or easily demonstrated methods of value. These may be your own solutions or in the instance of an AI platform even solutions developed by others.

2.4 Prevalence and Perception - the Goldilocks Principle

The Goldilocks Principle describes humans' natural inclination to find optimized solutions or settings pursuant to one's circumstances. In other words, the typical answers to preference lie in the middle of extremes. This is why Goldilocks, from the children's story, preferred the porridge that was neither too hot nor cold, but 'just right' in the middle. One can expect to find a similar experience when it comes to the task of balancing everything one may be able to produce with an AI solution vs. what is most meaningful for one's target audience. In order to balance this need, a clear understanding of prevalence related to one's use case is a must.

The impact and influence of disease prevalence for the use case cannot be over-emphasized. How often a set of circumstances present themselves in which a given AI model may act through a solution will have a significant influence on any end users' perception of both the quality and the value of that solution. For instance, if a disease is of very low prevalence, there is the risk users might forget the solution exists or may not recall how to use/interact with it.

An additional contributor is also the sheer number of applications and workflows clinicians (i.e. end-users) are already managing within their scope and practice. In this circumstance of low use, the value of an AI solution might also be perceived as low because it is used too infrequently. Even more alarmingly, low prevalence solutions may result in scenarios such as this:

- A disease with a prevalence of 3% and the single time the algorithm makes an error is on the one of the 3/100 instances on which that data is present, the perception of the end user is dramatically altered due to the error rate being 33%.
- Clinicians experiencing an unusably high error rate destroy the confidence and trust possible from a solution, its source, and any of the work done to deliver that AI feedback where it needed to be seen even though the disease prevalence is low.

On the opposite end of the spectrum, if a disease is highly prevalent and presented via alerts, users may be overwhelmed with the amount of notifications and stop paying attention, perceiving it as “too much noise”. While an algorithm with an accuracy of 99% when applied to a pathology with a 20% prevalence of what that algorithm is trained to find could be perceived as being extremely accurate and, therefore, helpful/valuable to the applicable patients and end users.

Therefore, when contemplating a clinical AI project, one must first understand the prevalence of what is being sought in order to accurately understand the requirements related to an algorithm’s effectiveness as perceived by the end user. For use cases with low or high prevalence, one must think about unique workflow capabilities to overcome the Positive Predictive Value (PPV) issues and risk of reduced adoption. It is important to balance the level of interaction necessary to establish initial trust in the solution as it will be solidified on a combination of its perceived accuracy, timeliness, and ease-of-use pertaining to both clinicians and patients.

2.5. Accuracy, Sensitivity & Specificity

Once integrated and timely delivered, the key metrics of success for each AI solution are its accuracy, sensitivity, and specificity. Accuracy is defined as the overall correctness of a given solution. Specificity is related to accuracy in that it is a measure of how reliably negative instances are correctly labeled as such. Sensitivity is how often something is correctly labeled

as positive. In other words, sensitivity is how often something is found and specificity is how often something is correctly identified as not present. The former helps with correctly identifying what we are looking for and the second helps with confirming what we are looking for is not there.

Within a clinical setting, accuracy (the combination of the sensitivity and specificity) is what will help bring that correct case to a clinicians' attention in a timely manner: i.e. an acute case is triaged and escalated to a care team appropriately.

A fourth factor related to these three is Volume. There is a distinct connection between volume of data analyzed and accuracy. As an example, if there are two solutions with the same specificity level of 95%. One analyzes chest CT exams while the other chest XR exams. The raw amount of false positive cases per day would be much higher in the CXR solution due to the sheer volume of studies produced by that modality vs. that of the chest CTs. While this is statistically to be expected, factors like this can have a great impact on both user perception of accuracy as clinical value - both of which can create risk to solution adoption. This is another illustration of why an AI model is different from a solution (building on our prior section).

When designing and planning for a solution, one will need to balance the variables and weights within the algorithm in order to achieve the outcome desired by its end users. Trust in this solution will need to be achieved through a careful balance between sensitivity, specificity, accuracy and how they are weighed against the volume of applicable data.

2.6. Workflow and User Interface

Workflow designed around an AI solution is another factor which will have a great impact on its adoption and user-perception of its accuracy. As such, one will benefit from a concerted effort in designing an intuitive workflow as it is as important as measuring an AI solution for accuracy and value.

Healthcare organizations, clinicians and patients will have a broad range of preferences, workflows, practices and habits. For instance, some are not bothered with receiving a high number of alerts related to their patients. This is sometimes seen as a kind of safety net and would prefer to 'see everything' rather than be concerned about something for their patients

being overlooked. Still others will feel quite the opposite - wanting alerted to only the most extreme of the patients or cases applicable to their practice. Therefore the user workflow delivered by the interface must be built in a flexible way to be able to accommodate for the wide spectrum of potential user needs and preferences. Certainly, where it is applicable, this may be accomplished algorithmically through an automated triaging mechanism, delivering exactly the type of result desired by one's target audience. An example of this could be a solution which doesn't alert a surgeon on every single abnormal finding but rather just those which meet a threshold of their determination.

Another means one may consider to accomplish this is via the filters within the UI delivered. If that UI allowed for the full spectrum of user-based filters (from a high to low frequency of alerts), one may be able to account for those who prefer to see all cases vs. those who only wish to interact with de novo findings. A simple example of filtering used to hone the alerts seen by an end user is to limit them by site across a network or time of day.

Additionally, ensuring that the module's interaction with the provider is in the day-to-day workflow of said provider is paramount. If one's solution removes the provider outside of their routine workflow, the provider will lose focus on the patient being cared for, limiting further the time allotted to care for the patient. If the workflow isn't perceived as easy and a seamless part of the provider's day, adoption will wane, and the solution will become obsolete.

In any case, the ability to deliver the preferred frequency of interaction will be directly related to the prevalence seen and the performance appreciated in one's solution.

2.7. Defensibility of Results

It is crucial to keep in mind that the introduction of a new tool into a clinicians' existing practice is bound to challenge their workflow and routine, which have been shaped following years of training and practice. A healthy sense of caution, therefore, is expected and should be welcome, given the immense clinical responsibility and duty for people's lives under their care. An important pillar for building trust in AI solutions and successfully introducing the anticipated change in everyday clinical workflow, is presenting the AI-generated data and model output(s) such that source information can also be efficiently reviewed. In a kind of side by side or quick reference manner, if an end user can see what information on which

an AI derived finding is based, their consumption and acceptance of that finding can be accelerated. Analytics, which we'll touch on again later in this document, is another manner by which the performance of a solution can be inspected, assessed, supported, and/or rejected. These analyses come in several forms which will be addressed in the MVP section of the document. The goal of any successfully designed and implemented AI solution is to relieve some of the clinical workload off of clinicians shoulders, thereby optimizing both their wellbeing, as well as patients' outcomes.

2.8. Analytics for Each Use Case

Rendering of a "Command Center Dashboard" by Gemini Advanced 2.0 Flash

The general purpose of clinical AI is to answer a question or set of questions. This produces net new information, and engagement with that information is what allows workflows, processes (clinical, operational, quality improvement, research) and outcomes to be impacted. All of that data and their impact should be continuously monitored and measured to ensure effective implementation and ultimately optimal outcomes (clinical, operational,

research, quality improvement). Some examples of such performance metrics are:

- The timing of each event involved in the workflow related to what triggers the analysis to begin, how long the analysis takes, and when a user has access to the output of the analysis
- Individual, site, and/or service line adoption
- Clinical metrics, i.e. clinical outcomes (e.g. mortality, morbidity, change in medications, new test/procedure, new clinic visit, visit to the ER, hospitalization, quality of life scores, patient-reported outcomes).
- Clinician well-being, i.e. ease of use, ease of interaction, ease of providing feedback to the algorithm, time it takes to perform certain tasks, ease of communicating with other care-team members and/or patients, alert-fatigue, and others.
- Operational effectiveness, i.e. reduction of rework completed by the providers or supporting staff.

Therefore, implementing clinical AI will require meticulous analyses for each use case in order to understand and, perhaps much more importantly, be accountable for the implementation of any new tool (or set of tools).

2.9. Adoption Consideration

Many of the topics and recommendations covered thus far in this paper are intended to help ensure the long term viability, sustainability, and successful deployment of an AI solution deployed into a production, clinical practice. While topics like: minimum viable product environment needed, regulatory requirements, costs, workflow, clinical impact, drift mitigation, unified visualization, agentic automation - several of the components which comprise a full solution rather than just a model - are all important, driving long term adoption of burgeoning technologies like clinical AI will require still further emphasis on the people who propose to leverage the above litany of technical characteristics. To address the users (clinicians) of this new brand of clinical tool so its potential is fulfilled, concentration on the following may be helpful to achieve broad, lasting adoption.

2.9.1. Enhance Clinical AI Education

As may be clear from the number of topics briefly covered in this document, there are many components of a successful AI solution when used in a production environment. While it is reasonable to suggest, clinicians, members of healthcare leadership, and IT are aware that AI tools are being produced at a blistering pace, they may not fully appreciate how they function. Though these same/similar applications are on our phones, computers, and trying to get into clinical applications everywhere, this does not convey understanding, just familiarity. It is actually now difficult to attend a healthcare industry trade show and find an IT vendor which is not promoting their ability to leverage AI in some manner. However, as of this writing, it may not be reasonable to expect today's clinical leaders and their teams to be well versed in the aspects covered in this document. Actually, that is one of our main points of focus. That is, we are attempting to point out the number of factors directly related to the success or failure of the implementation of AI into clinical practice. They must be well understood in order to avoid making mistakes related to a lack of awareness. Due to their complexity, power, expense, and potential for both positive and negative impacts, it would be good to advocate for broad elevation and enhancement of one's teams' understanding of Clinical AI literacy. With a more consistent educational base broadly experienced throughout an organization, one can expect to make higher quality decisions with respect to what kinds of tools to acquire or even develop.

While this could be a significant, longer term effort as health systems often employ hundreds or thousands of staff members who could benefit from this education, one positive aspect is there is an abundance of tools, many of them free, which could be leveraged. Here are some examples:

PCOM Library: **Artificial Intelligence in Medicine - AI Online Training**
Cousera: **AI for Everyone**
Harvard: **Leading AI Innovation in Health Care**

By investing in the education needed to help bring up the level of understanding of what AI is, how it functions, and what is required to allow it to flourish in a production healthcare environment, the quality of the assessment, procurement, deployment, and adoption of these new, clinical tools should only be improved.

2.9.2. Interdisciplinary Teams

There can be many stake holders relevant to even a single project leveraging one clinical AI solution within a single service line. Consider the following as relevant potential stakeholders for just a single use case which is designed to address an acute clinical pathology: service line leadership, senior IT strategy leadership, quality leadership, legal, AI governance, regulatory affairs, data and analytics officers, IT security, IT clinical applications, finance, and innovation.

It is these team members who often make up assessment, implementation, ownership, and/or oversight teams and committees for clinical AI projects. It is critical that engagement from these team members be made early and maintained as the project moves through phases of assessment, evaluation, validation, implementation, adoption, and ongoing utilization. Their combination of perspectives, understanding, and insight are often what is needed to ensure a successful outcome for each project.

2.9.3. Transparent Data Access

A significant factor in the ongoing reinforcement of a successful implementation will be related to how well and how clearly its users understand the impact the solution is having on their work and/or their patients. In as close to real time as possible, it will be important to have a strategy for providing this information to the applicable stakeholders, particularly those in enterprise and/or service line leadership.

The 'data' in question will come in a variety of forms, each with its own point of view to present. The following list is intended to give one an idea of how broad and ranging this information can be. It is very possible that one will have to pull from a correspondingly diverse set of sources in order to fully visualize and understand the impact from solutions deployed.

Accuracy Data - sensitivity, specificity, accuracy, areas under the curve, gold standard comparisons, etc. will all remain important metrics to continue to track throughout the lifetime of a solution. They may not require a real time understanding in each case. However, periodic analysis will be appropriate.

Clinical Outcome Data - this is literally anything related to the outcome the patient experiences in relation to an AI finding: Mortality, Morbidity, Length of Stay, treatment, admissions, throughput, time to treatment, error rates, complication rates, disease progression or remedy, and/or any reconciliation process.

Process Oriented Data - here one may want to focus on the efficiency/inefficiency achieved by implementing a novel AI solution. Examples of this type of analysis are: frequency of utilization, departmental/individual engagement, duration of use, individual feature utilization, deviations from intended use, and/or demographic characteristics of one's users.

Qualitative Data - as the saying goes, perception is reality. Therefore, one will want to be mindful of experiential information you may be able to reflect and/or quantify, like: surveys/questionnaires, direct feedback from within the given tool/solution, interviews among focus groups of users, and/or analyzing any communications among users which may reflect their engagement or satisfaction level.

Cost Effectiveness - like any other program or solution, maintaining an understanding of its cost/benefit ratio will continue to be critical. As was mentioned earlier in this document, the outcome/impact achieved by the AI solution must outweigh the cumulative effort taken to create, implement, and maintain that solution. The added variable for the clinical domain is the impact to patient care. Depending on the clinical effect which may be measured, one may find this to be a sufficient 'return' to justify the AI solution.

2.9.4. Change Management Culture

*Image of the famous scene from **Cool Hand Luke**: "What we've got here is failure to communicate." Luke was struggling under the pressures of "change management." ;)*

Changing IT systems has its challenges. But changing peoples' daily functions presents a new set of obstacles. BRIDGE is essentially advocating for both - changes in IT infrastructure and clinical workflows. Collectively we should change our expectations of what clinical informatics is now able to achieve. Because of this, a compelling opportunity to advance patient care is immediately before us all. However, as we have attempted to thoroughly outline within this document, there are many changes, both with the technology and workflow which will compel additional adjustments in how clinicians work or behave as a result. This is something one will, no doubt, want to manage carefully so as to help ensure successful initial implementation and longer term engagement. Here is a sample (each organization will want to make additions or deletions from the following to suit their preference and project) of some of the factors important for change management efforts related to clinical AI solutions:

- **Communication** - How often, in what manner, and on which topics teams communicate with each other will likely remain the most important change management factor. It is through good, clear communication in which the strategic vision or priorities for a project are established, leadership roles are selected, outcome expectations are described, desired workflow is outlined, teams are organized into their respective tasks, implementation is scheduled and executed, and all follow up and follow through tasks are assembled and completed. The number and variety of tasks involved in the creation and fulfillment of an idea that becomes a product will require excellent and repeatable standards of communication. This will have to be something which is maintained from the very beginnings, at which a governance structure is organized, all the way through the process by which a solution is adopted or rejected by its constituents. In order to keep this wide variety of necessary team members moving in uniform direction, transparency of this communication, like that which is required of the solution in question, will also be something one will benefit from providing.

- **Clinical Engagement** - BRIDGE is exclusively discussing solutions targeted to clinical workflows. It should not be surprising that there will need to be the appropriate clinical engagement from the beginning, middle, and end of each of these solution projects. It is also noteworthy that these projects don't actually end. As we have discussed, the need for ongoing drift mitigation, monitoring, and correction will require a consistent level of energy and emphasis throughout a solution's life cycle. Any adjustments which need to be made along the way should be checked against the clinical leaders associated with that solution in order to ensure new adjustments being made are not detrimental to patient care. Therefore, maintaining a means of continuous clinical consultation will also need to be accommodated.

- **Training and Implementation** - When an individual clinician on a team of providers is being trained for the use of a new tool, this will often be the first time he/she will have a true opportunity to grasp exactly how this tool will affect how they work. It is, therefore, critical to make a good first impression. Emphasis on thorough and clear communication and demonstration of what each solution is designed to accomplish will pay significant dividends for the proposed project. It is at this point that consideration will need to be given to what documentation is also provided to an end user as well as how feedback is collected from that individual. Implementations come in a variety of forms - enterprise wide down to the individual contributor. With respect to clinical AI solutions, it is likely one will want to start with a collection of keenly interested, clinical champions who will embrace the need to take a creative and iterative approach to a problem solving style of tasks that is the implementation of net new clinical infrastructure and tools. Establishing feedback loops from these champions, considering multiple modes of training materials and media can be important, documenting and sharing any quick wins established by this team, as well as following through with a steady, regular cadence of regrouping as a team to share what is and is not working as expected will all be critical components of a successful change management strategy.

- **Culture** - At the risk of burying the lead, all of the above comments in this section can be summed up as the culture of innovation needed to create and then follow through with all that is needed to successfully deploy novel AI solutions into a clinical space. In so doing, it is appropriate to say that the deployment of clinical AI solutions isn't something that one does, as if it was just an item on a check list of tasks for the day. Rather, it is something which needs to become a component of your institutional characteristics. Simply: it's what you are. This cultural adoption becomes part of you as you follow through with what clinical AI needs to be done well and for each of the patients involved to be positively affected.

Given the emphasis on a culture of change, emphasis should appropriately be placed on the tactics and practices which increase its probability. Therefore, successful CM practices should be founded on methods which put it into practice. Focused alignment of both individual and organizational goals, minimizing disruption and ensuring long-term success, will need to be a point of emphasis throughout any clinical AI transition process.

An area in which change management will not be effective if it is viewed as a top-down mandate, is simply a PDF of rules linked on the intranet, or is an adjustment made simply for monetary purposes. Given this, here are just a few change management tactics which may help establish one's culture of change:

- **Baseline Your Current Process:** Knowing and quantifying your existing process will allow you to better understand where you need to go.
- **Allow Engagement:** Involve clinicians, administrators and patients in shaping AI initiatives at key stages to build a shared vision. Don't shy away from open communication.
- **Empower Champions:** Leverage their expertise and connections when drafting policies and procedures that will impact end-users.
- **Encourage Communication:** Promote open dialogue about AI – what it does, how it will be used and safeguards in place – to build trust.
- **Document decisions:** Ensure that each precedent-setting decision is documented as well as the rationale for the decision.
- **Be Transparent:** Be clear on how AI outcomes will be measured and usage will be monitored. Put this information in a place that is easily accessible to those that need to know.
- **Ongoing Education:** Emphasize ongoing training aligned to metrics so skill and knowledge gaps are addressed.
- **Celebrate Wins:** Regularly share the impact of AI on patient outcomes, staff satisfaction and organizational performance.

For a deeper dive into how other organizations have sought to establish trusted practices for AI implementations, please see this **summary from 14 different organizations** provided by Gemini Advanced v1.5 Pro Deep Research.

3. Development and Technical Foundations

3.1. Minimum Viable Components of a Production AI Use Case

Once one understands each of the above factors and has produced what is believed to be a sufficient AI-driven solution ready for production use, the next immediate need is to have a technical mechanism sophisticated enough to implement that solution while balancing all the criteria which has been outlined. Below is a list of the minimum requirements for such a method or system. This list is not meant to be exhaustive. Rather, it is intended to outline the minimum of what one may wish to consider necessary for the longer term success of one's solution.

3.1.1. Agentic Mechanisms

In order to operate at scale with the necessary trust mentioned earlier in the document, a production AI solution will need to run automatically within a heterogeneous environment with respect to data presentation. Even standards mentioned earlier (like DICOM or HL7) have many vendor and deployment specific variations for which it can be difficult to account in the absence of an automated mechanism to do so. This is precisely why Agentic forms of AI have been created so the needed tasks of a high pressure, fast paced, diverse environment can still be automated in a trustworthy manner. There are at least three different mechanisms which will need to be done in an automated, agentic fashion in order to be successful over the long term (beyond initial deployment).

3.1.2 Normalization

It has been mentioned that data heterogeneity is a significant factor which must be overcome in order to reliably deploy an AI solution. This is due to the unpredictable nature of healthcare data. Within an imaging study, edge of film anatomy is routinely captured which may contain relevant findings applicable to an imaging algorithm. A language based tool may be configured to search, find, and summarize a wide variety of data elements which could be produced, documented, edited, dictated, printed, etc. throughout the EMR (each deployment of which is unique). For workflows which require the use of both imaging and non-imaging data, the presentation of that data gets even more varied. Hence, the need for a

mechanism to normalize the nature in which data is presented associated with and within an imaging study, report, document, set of discrete fields, etc. is a prerequisite in order for a set of algorithms to be able to predictably take action on the data presented. An absence of this normalization mechanism results in the inconsistent application of the deployed solutions which have been designed to make use of that data. Inconsistency will lead to distrust. Distrust to disuse and the ruining of a project. Therefore, one is strongly encouraged to ensure the sufficient normalization mechanism is in place so the likelihood of their project's success may be improved.

3.1.3. Activation

Once a consistent presentation of data is available for an AI solution, automated means of activating that solution needs to be available. Please keep in mind that, as mentioned before, it is likely that any novel AI solution will need to be activated within the context of existing solutions without disrupting their function and, if done optimally, could even augment the quality of those solutions. The intent of new solutions is to make things better rather than serve as a distraction, slow users down, and/or simply give them more data to adjudicate. In order to deploy such solutions successfully, they must be done in a manner that does not add any additional work to the personnel (of which there is a shortage) involved in the use case(s). Simply, it must be automated.

3.1.4. Mitigation

Just as it requires significant data presentation and analysis to create an AI-driven solution in the first place, it requires further, automated analysis to ensure that it behaves as it has been designed over time. This is because all algorithmic solutions drift. Meaning, they eventually stop working with the original level of quality and output produced upon original deployment. Drift can present itself in two basic forms: Interveniable and Inevitable Drift. Interveniable Drift simply refers to a degradation in performance of an algorithm over which a user or organization can have some influence. Meaning, it can be remedied due to the cause being related to a factor over which one has controls. For example, if an imaging algorithm is designed to find small, difficult to see pathologies on CT, it will very likely require access to the 'thins' portion of how a CT scan is produced. If anything, like an update to a scanner, disrupts its access to this fidelity of data, the likelihood of it functioning properly goes down. If this is discovered to be so, an alteration in the algorithm's access to that data fidelity can be provided, and one should expect its performance to return to within expected limits. However, if one is using or deploying Federally regulated algorithms, those are not allowed to be updated to an individual site's data and will be subject to Inevitable Drift. This is because it is unable to learn from the site's heterogeneity (variables related to its modality and patient mix specific to its community). Over time the algorithm will be exposed to more

and more data combinations and presentations on which it was not trained, degrading the quality of its output. Once that output is no longer within expected ranges, it is likely that it will need retraining and resubmitted to the appropriate regulatory body. Therefore, no matter the type of AI solution, implementing and maintaining a method by which the appropriate proxy information can be compared to live algorithm performance will be an additional prerequisite in order to ensure that a solution can be deployed, work as expected, and continue to meet that expectation over time.

3.1.5. Implications of Agentic Mechanisms

Agenticism is what will allow one's solution(s) to be implemented at scale and alongside many of other solutions. This has been labeled as a prerequisite in nature because of the narrow margins (labor, time, finances) which beset the whole of global healthcare today. Simply put, healthcare is in a state in which it needs to be able to do more with less. By leveraging Agentic, automatic means of releasing the full power and capability of artificial intelligence, one can reasonably expect to make a positive contribution to his/her organization in doing so.

3.1.6. Enterprise Integrations

As has been mentioned throughout this document, an AI solution may need simultaneous access to a variety of systems and data sources on which it is dependent in order to reach its goal of being adopted. At a minimum, it must have constant and consistent connection to the needed Agentic Mechanisms (mentioned immediately above), appropriate data sources upon which it can leverage its analysis, native workflow solutions within which its end users already work, as well as adjacent systems which control identity management of those users as well as when those users work. If just one of those connection points (integrations) are removed from an AI solution, its ability to perform as designed (achieve adoption) can be irrevocably diminished.

3.1.7. Workflow Flexibility

One of the wonderful opportunities for AI solutions to contribute to is the chance for health practitioners and systems to contemplate changing what their current approach to providing healthcare to be. Today's environment is typically built of many layers of manual processes within each medical service line. However, a patient, especially acute one's, may have to rapidly be exposed and subject to a variety of them very quickly. Rather than having merely more service line dependent automations (not that those are negative), AI solutions have the opportunity to automate how and to whom data is disseminated and presented. This ability can provide the means for service lines to lower the silos built around them so they can more effectively collaborate with their colleagues when 'sharing' a patient.

Doing this at scale within health systems will require a high degree of workflow flexibility and configuration such that one's end user audience can design the care pathway they believe is most appropriate for the automation being delivered. Afterall, with AI, one will be introducing something brand new. One should expect to provide iterative flexibility so users can gain experience with a new tool, learn from that experience, and then collaborate to help optimize the impact from a given tool. This will require functions like controlling who has access to the tool, when, and at which endpoint. It will need one to plan for means of piloting to a smaller group of users while also providing the means of scaling up to many hundreds and/or thousands of users, depending on the size of one's organization and prevalence of the use case. If one is delivering a use case which enables disparate service lines to collaborate in real time, there will almost certainly need to be a means of synchronization of AI-derived information across disparate endpoints. These are just a few examples to help provide the reading with a framework of the types of workflow iterations one can expect to be required to provide.

3.1.8. Synthesis of results

Today the main means of synthesizing disparate types of data across the healthcare landscape is the clinicians themselves. As mentioned earlier, the shortage of these resources is what is spurring the very innovation of AI-driven solutions. However, due to the number of them being produced, there is the very common need to allow for a variety of them to contribute at the same time, in real time, to a single patient's study, encounter, etc. This presents one with the subsequent challenge of ensuring there are sufficient means in place to make a unified presentation of the contribution of a variety of tools to the clinician in a manner that satisfies the needs already expressed earlier in this document.

For example, if there were a language based solution deployed which accurately and reliably synthesized a patient's history in a user friendly and helpful format, there isn't a single kind of physician who wouldn't benefit from always having it available to him or her. So where does one display such data? Because of the many specializations across healthcare, there can be a correspondingly variety of disparate means within which this summary information can be expected to need to be displayed: RIS, CVIS, EMR generally, patient facing summaries, interventionalist workflow applications, and many others. If one then took the next logical step to take the summary information and provide risk stratification for that patient, depending on the answer to that stratification, a secondary subset of users will likely need to readily consume and perhaps even be alerted to this information. When they receive that alert, they'll need to have their own subspecialized synopsis of multimodal information in order to efficiently render (synthesize) decisions regarding patient care. The further one follows this logic, the more complex the combinations of synthesizing a variety of AI driven outputs can get. Therefore, thorough thinking through these real world clinical scenarios is a must when designing a novel solution.

3.1.9. Variety of User Interfaces

Building on the point made immediately above, one will need to ensure that a minimum variety of user interfaces are available to be leveraged within a novel workflow. Some examples of those are: embedding within the EMR (requires EMR integration, likely bidirectional), EMR overlay, PACS worklist, PACS desktop, mobile user interface, and/or web portal. Depending on the solution designed any combination of these may be applicable to the respective audience.

3.1.10. Modular Visualization and Storage

One should note this factor, as of this publication, primarily involves imaging related use cases because the results of which must be seen, rather than read as in the case of a language based solution. When deploying imaging based solutions, as was mentioned earlier, those will have a number of different types of regulatory approval which are then mapped to different visualization and storage policies within one's organization. Therefore, a mechanism to comply with these medicolegal policies will be a must for production use. For instance, on a single chest CT, one could have a CADt solution for triaging hyper acute findings, CADe for the diagnosis of present pathology, and a measurement algorithm to standardize data capture and automate report production related to that data. Each of those will likely be visualized differently from one another: CADt via an overlay application to keep it out of the long term record, CADe within the diagnostic record, and the measurements reviewed within the appropriate reporting application to ensure the quality of that documentation. This kind of scenario will play out over and over again in an ever increasing amount of complexity as more and more solutions are made available either via commercial or of one’s development.

3.1.11. Security and Governance Needs

To fully realize the benefits which are possible from AI solutions which are implemented well, health systems must mitigate certain risks. Safeguarding trust, safety and ethics are key, and achieving this requires robust internal governance: a framework of policies and oversight mechanisms that ensure AI is implemented responsibly, transparently and safely. Just as the digitization of diagnostic and electronic health records (EHR) revolutionized care delivery, the widespread adoption of clinical AI demands a careful balance between innovation and caution. This isn’t just about adopting AI; it’s about doing it in a manner which engenders a process of compounding trust (consistent with the velocity of trust outlined in Section 3). AI adoption demands a holistic perspective. Governance and cybersecurity aren’t isolated tasks

but interconnected pillars of a successful, scalable strategy. By weaving these principles into the fabric of organizational operations, health systems can unlock the full potential of clinical AI while safeguarding the trust and safety of their communities.

AI systems influence diagnoses, workflows and patient outcomes, yet their inner workings can feel impenetrable – a black box, so to speak. Transparency changes that. It opens the box, lays everything on the table, and says to every stakeholder: "This is how it was created, this is how it works, and this is why you can trust it." By emphasizing transparency, health systems show patients and clinicians that technology can have a meaningful and responsible impact. Ultimately, transparency is more than a practice; it's a promise a promise to put people first in every AI-driven decision.

3.1.11.1. Building Governance That Works in the Real World

Effective governance frameworks are more than policies. They're living, breathing structures that adapt to the dynamic nature of healthcare, regulations, and technology. Here's how health systems can build governance frameworks that succeed:

Integrate AI Expertise into Existing Structures - AI governance doesn't have to reinvent the wheel. Embed internal and external AI knowledge into existing clinical and operational committees to ensure decisions are informed by expertise without adding unnecessary bureaucracy.

Foster Multidisciplinary Collaboration - The best governance frameworks are collaborative. Bring together executive champions, clinical leaders, IT experts and frontline staff to create coalitions that reflect diverse perspectives. This ensures governance decisions are practical, inclusive and aligned with organizational goals.

Prioritize Transparency - As we discussed earlier, transparency is the foundation of building trust in AI. By demystifying AI systems, governance frameworks can dispel the black box perception and build confidence.

Commit to Continuous Improvement - Governance isn't static. Regularly assess frameworks to ensure they evolve with new AI use cases, regulatory updates and organizational priorities. This adaptability keeps governance relevant and effective.

3.1.11.2. Best Practices for Building AI Cybersecurity

Cybersecurity isn't a wall in AI governance; it's part of the foundation. Without it, the benefits of AI are overshadowed by vulnerabilities that could compromise patient safety, violate compliance standards and erode confidence in the technology. Cybersecurity in clinical AI requires more than firewalls and encryption. It demands a proactive, multi-layered approach that anticipates threats before they arise. From data breaches to adversarial attacks, the risks evolve as rapidly as the technology.

Here's how health systems can build and maintain effective cybersecurity:

- Conduct Regular Risk Assessments
- Adopt Privacy-by-Design Principles
- Forge Collaborative Developer Relationships
- Stay Ahead with Updated Protocols

In addition, how AI is delivered into a health system directly impacts its security. Point solutions, marketplaces, and platforms each have unique implications:

Point Solutions are a boutique deployment focused on specific tasks that can create data silos and offer limited scalability. They will each have their own requirements and/or abilities to provide the necessary security functions. Therefore, assessment, implementation, and maintenance of those parameters must be repeated for each point solution implemented.

Marketplaces offer a hub for discovering and choosing different AI solutions with the primary incentive of adoption being a simplified contracting process. The deployment may be centralized or discrete installs of each solution. Solution deployment and training will oftentimes be handled by the algorithm supplying vendor and not the marketplace vendor. Marketplaces often require health system expertise to navigate and may experience varying quality and system compatibility. From a security perspective, unless the marketplace provider is unifying these processes, the same approach outlined for point solutions must be repeated.

Platforms can provide a centralized infrastructure for scalable AI implementation with robust integration to in situ systems. One will want to fully understand the extent to which the security functions applied to the platform are extended to each AI solution it deployed. If applied equally, the initial implementation, which can require an upfront investment of resources, but could also pay dividends as a unified security policy and structure are uniformly applied to each solution.

Of these, platforms often provide the most secure approach, minimizing silos and enabling centralized oversight of compliance and cybersecurity measures. Close and early consultation with one's security team in order to completely understand their needs will help ensure this component of the creation process is well executed.

3.1.11.3. Real-World Cybersecurity Scenarios in AI

Cybersecurity in clinical AI requires more than firewalls and encryption. It demands a proactive, multi-layered approach that anticipates threats before they arise. From data breaches to adversarial attacks, the risks evolve as rapidly as the technology itself, making cybersecurity a dynamic and essential component of AI adoption.

Below are real-world scenarios demonstrating both threats and successful mitigation strategies to clinical AI safety.

Adversarial Attacks in Medical Imaging

Scenario: Image Manipulation in Radiology
Researchers have shown that adversarial attacks can subtly alter medical images, such as CT scans, to create or remove signs of disease. In one test, AI systems misdiagnosed manipulated images with 99% confidence – despite the changes being imperceptible to human radiologists.

Mitigation Strategies

- Incorporate adversarial training to expose AI models to manipulated examples and improve resilience.
- Use encryption and blockchain to ensure image integrity throughout the workflow.
- Deploy anomaly detection systems to flag suspicious patterns in medical image data.

Privacy-by-Design in AI Deployment

AI systems must also safeguard patient data while ensuring compliance with regulations like HIPAA and GDPR.

Success Example: Federated Learning in COVID-19 Research
During the COVID-19 pandemic, hospitals used federated learning to train AI models without sharing sensitive patient data. This approach allowed institutions to collaborate while maintaining compliance and accelerating research.

Key Success Factors

- Decentralized data architecture to avoid centralizing sensitive information.
- Strong encryption to protect shared model parameters.
- Transparent collaboration agreements to ensure regulatory compliance.

3.2. Hallmarks of a Good Solution

3.2.1. Reproducibility

When contemplating building a solution, one of the earliest decisions which needs to be made is the intent of the scale for that solution. Meaning, does one want to try and create something which could be leveraged by others elsewhere or focus on usage within one's own institution? If it is the latter, the project will be less complex.

Leveraging an internal solution developed within one's own site involves pursuing an IRB (see governance section above) to cover the use of that solution. If the aspirations of a given AI project are to enable the solution to be leveraged outside of one's facility, then FDA clearance or CE Mark (or equivalent for the given market) will be a necessary path down which one will need to proceed.

Given there are 950 FDA cleared clinical AI solutions (as of August 2024) for medical imaging alone, the road to regulatory clearance/approval is a well worn one, though it has mostly been paved by commercial vendors. As is mentioned above, there are a number of different regulatory processes and agencies involved in this space. Once the framework of regulatory approval for a given solution is well understood, there are significant advantages which can be achieved with a solution(s):

- Portability - a model you create could be leveraged by others
- Greater impact on medical care to every country and community
- Commercial opportunity for additional revenue for your organization

While those are each laudable, in order to prepare a solution to work well enough to function equally within disparate organizations (portability, reproducibility), there is a significant burden on the data on which a solution is trained. Not only does bias need to be kept as low as possible, the model created needs to be able to work equally in heterogeneous environments. This heterogeneity comes from four main sources:

1. Patients with differing backgrounds, social determinants, and ethnic/genetic makeup
2. Wide variety of combinations of modalities, study parameters, and data cleanliness
3. Site level preferences and variation in the performance of diagnostic tests
4. Provider level variation in documentation techniques

The heterogeneity of the environment (patients, modalities & data presentation) in which a solution will be deployed will be touched on again within the MVP section later in the framework. The challenges presented by data, modality, and patient heterogeneity are such that they will remain something one must account for within the application of any AI-driven product for the foreseeable future independent of whether or not a solution is used internally or externally.

Further investigation into concepts related to overcoming data heterogeneity related challenges **can be found here**.

3.2.2. Clinical AI's Opportunity to Define Resilience Requirements

Artificial intelligence (AI) offers unique opportunities in healthcare, elevating many applications into the "mission-critical" category. Clearly identifying the uniqueness of data and the criticality of system access provides a framework for approaching resilience and redundancy in healthcare settings.

3.2.3. Centralized Implementation and Categorization Considerations

While a centralized platform simplifies many aspects of AI implementation, categorizing use cases remains essential. Different operating parameters or requirements will influence the implementation process, even within a centralized approach. Centralized systems may also be able to offer lower complexity and less costly approaches.

3.2.4. Resilience tiers

Healthcare systems commonly use a tiered categorization framework to prioritize systems and applications. This framework typically employs a descending numerical order, from most essential (Tier 0 or Tier 1) to least essential (Tier 2 or Tier 3):

- **Tier 0: Mission Critical**
  - Reserved for systems and information essential for delivering timely healthcare consistent with the standard of care.
  - Examples include Electronic Health Records (EHR), Picture Archiving and Communication Systems (PACS), and patient monitoring systems. These require continuous operation under all circumstances.

- **Tier 1: Mission Essential**
  - Systems vital to functionality and operations but tolerable for short downtimes.
  - Examples include non-critical documentation, scheduling, and billing systems.

- **Tier 2: Mission Enhancing**
  - Non-essential systems that improve care but whose absence does not impede it significantly.
  - Examples include predictive early warning systems, research-related informational systems, or analyses that can reasonably be performed manually by staff.

3.2.5. Importance of Accurate Categorization

While it might be tempting to label all systems as "mission critical," overcomplicating categorization can hinder, rather than help, long-term operations. Therefore, when determining the appropriate tier for a new AI use case or system, consider the following factors:

1. **Uniqueness**
   - Is the information or functionality novel, or can it be replicated through manual processes?
   - Example: Automated analysis of CT perfusion acquisitions for stroke care, which provides critical insights not easily derived from raw data.

2. **Fit Within Care Delivery**
   - How well does the system align with existing care workflows?
   - Example: CT perfusion for stroke care has become part of the standard of care following the DAWN and DEFUSE 3 trials, but its necessity depends on clinical scenarios, such as the timing of symptom onset.

3. **Timeliness Requirements**
   - Are the results needed immediately, or can they be delayed without impacting care?
   - Example: Outpatient exams can reveal an unexpected critical finding. These results may need to be expedited in order to prevent the patient from leaving prior to receiving care.

4. **Potential for Reputational Harm**
    - Could system failures erode trust among clinical teams or harm patients?
    - Proactive monitoring, clear policies, and immediate responses to concerns are essential to maintaining trust and mitigating reputational risks.
    - Example: **JAMA research study** referring to a poor performance sepsis model "widespread adoption despite poor performance raises fundamental concerns about sepsis management on a national level."(5)

3.2.6. Dynamic Nature of Categorization

Categorization is not static and must adapt to changes in clinical and administrative priorities. For instance, as new clinical trials and/or solutions emerge, their findings may affect the categorization of AI use cases, such as CT perfusion in stroke workflows, which evolved from an academic project at Stanford and has since become regarded as a standard of care for stroke evaluation.

Categorization can also evolve over time as workflows and adoption rates change. For example, an initially Tier 1 application may become Tier 0 after robust adoption, especially in acute care pathways where time-sensitive decisions directly impact patient outcomes. For instance, when speech recognition systems first became prevalent in use, traditional dictation/transcription methods were leveraged in times of any failures. However, as those speech driven products became regarded as the new standard, use of legacy systems as a fail over was no longer regarded as adequate. Hence, the speech driven systems of today are deployed with their own failover strategies so that the standard can continue in the event of a failure to the primary system.

3.2.6.1. Aligning Categorization with Institutional Processes

Once a use case is categorized, aligning it with existing institutional protocols becomes paramount. Most healthcare facilities measure system availability using the "nines" of uptime. For example, a Tier 0 mission-critical system with four nines of uptime (99.99% availability) allows for approximately 4-5 minutes of downtime per month or 52 minutes annually. Achieving this often requires:

- Redundant local systems (including networks),
- Multiple internet connections from different vendors, and
- Temporary downtime workflows.

One will likely want to establish similar tolerances and time calculations are warranted for the lower tiers of AI application criticality to a given department, service line, and/or enterprise. Particular attention will likely need to be made to whether or a given AI solution can be delivered in a manner which supports the items bulleted immediately above.

3.2.6.2. Failover Protocols for Critical Systems

For Tier 0 and Tier 1 systems, failover protocols are essential. When systems go down, there must be a clear plan for continuity, whether by leveraging alternate systems, engaging external services, or implementing manual workflows. Clear documentation, communication protocols, and awareness among stakeholders are critical to maintaining resilience during outages.

This new age of AI solutions will not save any of us from the consequences of overlooking or underestimating the impact of having a well thought out plan for redundancy and resiliency as these new tools become more and more commonly used for day to day clinical workflows and operations.

3.3. Industry and Development Frameworks

3.3.1. IHE Profiles

In addition to the standards that healthcare systems use to communicate, such as DICOM [1], HL7 [2], SNOMED [3], LOINC [4] and others, there's an organization, Integrating the Healthcare Enterprise (IHE) [5], that coordinates the development and validation of "profiles" to address interoperability workflows. It's not enough to have a standard that just covers 'what data is shared', and 'how the data is shared', but 'to what purpose is this data being shared'. This is what IHE covers.

IHE has developed a number of assets relating to AI in radiology. These profiles include:

- **AI Results (AIR) [6]:** This profile covers the structure of an AI result. Despite that there are standards available, there remain many ways for a segmentation result to be represented. It could be markup (GSPS), a secondary capture series, a bit mask, a heat map, a structured report, a FHIR Observation, or several other ways. AIR provides, for a given result type, the type of object that should represent it for all software applications to implement.
- **AI Workflow for Imaging (AIW-I) [7]:** This profile covers how AI services can be requested and serviced. It is simply not scalable for "PACS to send a full DICOM study to every AI end-point," and there is a need for an orchestrator actor to coordinate what jobs are requested and how they are being completed. AIW-I provides a mechanism in which this can be done.

- **AI Interoperability in Imaging White Paper [8]:** This is a background read of the broader picture for all AI components in a radiology and enterprise imaging workflow, enumerating many use cases, personas, systems, and workflows. While not a prescriptive workflow, it provides the basis for exploring what these actors do in order to shape how they get built into the future.
- **Imaging Diagnostic Report (IDR) [9]:** This profile describes how radiology reports can incorporate insights from many different sources, including AI results, into a unified report.
- **Prioritization of Workflists for Reporting (POWR) [01]:** This profile describes how insights can be used, in tandem with other rules, to structure the worklist and the relative priority that work is being done.

These profiles help inform how systems can be crafted and scaled at healthcare enterprises.

More information regarding the Integrating the Healthcare Enterprise **found here**.
IHE Radiology White Paper on AI Interoperability summary provided by Perplexity.

3.3.2. Development Frameworks

There are many frameworks, SDKs, APIs and other development tools that can be used to build AI solutions in healthcare. While not intending to be an exhaustive list, these are notable open source tools used in this space.

- **Project MONAI**: Medical Open Network for AI [1]. This project provides training (MONAI Core), labelling (MONAI Label), and deployment (MONAI Deploy) platforms for medical AI.
- **PyDicom** [2]: This SDK for Python provides the tooling necessary to work with medical imaging files, including the ability to read, write, and extract metadata and pixel payloads. This library is frequently used with frameworks like NumPy to work directly with the imaging data.
- **PyNetDicom** [3]: Related to PyDicom, PyNetDicom provides the functionality to communicate with medical imaging systems.
- **HighDicom** [4]: To craft more advanced DICOM objects used in the results and analytics of medical images, HighDicom provides utility in mapping the data into DICOM objects.

3.3.2.1. MONAI Project and Core Library

Project MONAI is the leading open-source toolkit for medical AI research and product development. It has been downloaded over 4.5M times, has had contributions from over 270 developers from around the world, has been featured in over 3,000 publications, and has been used to win over 18 international-conference and Kaggle medical imaging challenges. It has also been used to create multiple regulatory-approved medical products. It has been released under the permissive Apache 2.0 license which allows its free use in research and commercial applications.

In 2024, MONAI's impact and community engagement was greatly accelerated. It was expanded to include GenAI frameworks for synthesizing medical images and for vision-language models. It also introduced foundation models for synthetic CT generation (MAISI), 3D whole-body CT segmentation with 127 classes (VISTA-3D), and pathology segmentation (VISTA-2D). Project MONAI also established eight working groups that are open to public participation to help steer the future development of MONAI. Those working groups are detailed on the MONAI website and summarized below:

- **Evaluation and Benchmark Working Group** - accuracy and efficacy metrics
- **Education Working Group** - graduate-level coursework
- **Ophthalmology Working Group** - tools and foundational models
- **Outreach Working Group** - website and master slide deck
- **Deploy Working Group** - DICOM, FHIR, and clinical workflow support
- **Core Developers** - high-quality software processes and infrastructure support
- **Federated Learning Working Group** - learning across institutions and locations
- **Human-AI Interaction Working Group** - AI-assisted annotation across domains

MONAI Core is the central library of Project MONAI. It provides domain-specific capabilities for medical AI model research and development. Those capabilities span:

- **Medical Data I/O**: MONAI supports images in over 30 formats such as DICOM objects, NRRD, Nifty, MetaIO, and TIFF, with much of that support coming from integration with the Insight Toolkit (ITK, https://itk.org) and nibabel. It also includes support for ECG and other forms of medical data.
- **Medical Data Processing**: MONAI contains custom methods for processing radiology, pathology, and signal (e.g., ECG) data to prepare it for AI analysis. It also supports the integrative use of ITK, SimpleITK, and other medical and vision image processing libraries.

- **Cutting-edge AI Algorithms**: The open-source licensing and community engagement practices of MONAI has attracted contributions from the best and brightest in academia and industry. The authors of leading medical AI papers are encouraged to contribute their code to MONAI. This has led to MONAI containing the latest transformer, GenAI, and AutoML methods, for a diversity of use cases and modalities.
- **Federated Learning**: The algorithms of MONAI also address practical concerns. For example, MONAI can leverage Federated Learning frameworks that enable collaborative learning across institutions while maintaining data privacy.
- **Clinically-motivated Evaluation Metrics and Guidelines**: MONAI Core offers a wide range of evaluation metrics for assessing medical AI model performance. The Benchmarking Working Group collected these metrics and published guidelines on their use, and that publication was recognized by Nature Methods as their third most influential paper of 2024 (6).

MONAI Core and the other libraries of Project MONAI also follow high-quality software practices to facilitate reproducibility and speed the translation research to clinical practice. These software practices include rigorous standards for automated software testing and detailed documentation. Additionally over 50 tutorials that comprehensively cover the MONAI foundation are available, and multiple workshops are held annually at international conferences to provide expert-guided hands-on lessons and speed its adoption among the current and next generation of researchers and product developers.

3.3.2.2. MONAI Label

The quality and quantity of annotated data available for training or evaluating an AI algorithm is often the deciding factor in its accuracy and generalizability. In recognition of the critical role of annotated data in medical AI, Project MONAI created MONAI Label. This client-server library targets reducing the time required for data annotation by 70%. MONAI Label has been integrated with industry-standard medical imaging tools such as 3D Slicer and OHIF for radiological images, CVAT for video, and Digital Slide Archive and QuPath for pathology images. MONAI Label is an active learning framework that suggests initial annotations, learns from the edits to those suggestions made by experts, and intelligently selects new cases for annotation to strategically speed the learning process.

3.3.2.3. MONAI Bundles

Once a MONAI model has been created, it can be published as a MONAI Bundle. These bundles are intended to capture the pre- and post-processing steps as well as the AI algorithms essential to a medical AI workflow. In 2025, work began to refactor MONAI

Bundles to adhere to the Hugging Face pipeline standard. Adopting the Hugging Face standard allows the greatest portability of MONAI models for research.

There is also an initiative to capture details on the data and algorithms used to train the model in a MONAI Bundle, and thereby address AI model trustworthiness and responsible use. That initiative has been inspired by the CHAI initiative, which is discussed next.

3.3.3. The Coalition for Health AI - CHAI

The Coalition for Health AI (CHAI) is another organization gathering stakeholders from across the health sector to come up with consensus-driven definitions, considerations, metrics, and tooling to help developers and customers of AI adopt responsible AI principles.

CHAI defines a trustworthy and responsible AI solution as one that accounts for the following elements:

- Usefulness
- Safety
- Accountability and Transparency
- Explainability and Interpretability
- Fairness
- Security and Resiliency
- Privacy-enhanced

It is imperative that implementing organizations also evaluate AI solutions according to the same principles. CHAI created the Responsible AI Guides to aid organizations with specific considerations tied back to each of the core principles and aligns to the CHAI AI lifecycle. CHAI's AI lifecycle consists of the following stages:

1. Define Problem & Plan
2. Design the AI System (for developers)
3. Engineer the AI Solution (for developers)
4. Assess
5. Pilot
6. Deploy & Monitor

To operationalize the core principles, CHAI recommends the following considerations for the particular lifecycle stages of AI solutions:

Define:

- **Objective Setting**: Clearly articulate the problem the AI system aims to solve, ensuring alignment with clinical needs and organizational goals.

- **Stakeholder Engagement**: Involve diverse stakeholders, including patients, clinicians, and ethicists, to gather comprehensive requirements and perspectives.
- **Risk Assessment**: Identify potential ethical, legal, and social implications, considering factors like bias, privacy, and patient safety.

Design & Engineering:

- **Data Management**: Ensure data quality, representativeness, and compliance with privacy regulations.
- **Algorithm Design**: Incorporate fairness and bias mitigation strategies, and design for interpretability and transparency.
- **Documentation**: Maintain thorough records of design choices, data sources, and ethical considerations.

Validation:

- **Testing**: Conduct rigorous validation using diverse datasets to assess performance, safety, and fairness.
- **User Feedback**: Engage end-users in testing to evaluate usability and gather insights for improvement.
- **Regulatory Compliance**: Ensure the system meets all relevant legal and regulatory standards.

Deployment:

- **Implementation Planning**: Develop a clear deployment strategy, including training for users and integration into existing workflows.
- **Monitoring Framework**: Establish mechanisms for continuous performance monitoring, error reporting, and user feedback.
- **Transparency**: Provide clear information to users about the AI system's capabilities, limitations, and decision-making processes.

Monitor:

- **Performance Evaluation**: Regularly assess the AI system's outcomes to detect issues like performance drift or emerging biases.
- **User Support**: Maintain open channels for user feedback and provide timely assistance to address concerns.
- **Continuous Improvement**: Update the AI system as needed based on monitoring results and evolving best practices.

CHAI has formed working groups to apply these considerations to specific use cases that reflect the most common applications developed and sought after by customers. To reflect an important gap in current market needs, CHAI is compiling a list of methods, measures, and metrics associated with each use case to help validate specific AI models and solutions. For example, a sepsis model can measure performance with the following methods, measures, and metrics:

Method:

- Ground Theory Analysis

Measure:

- Sequential Organ Failure Assessment (SOFA) score
- Risk ratio (outcome rate if exposed to AI model over outcome rate if not exposed to AI model)

Metric:

- Area under the curve (AUC)—receiver operating characteristic (ROC) (AUC-ROC)
- Area under the curve (AUC)—precision recall curve (PRC) (AUC-PRC)
- Precision (or Positive Predictive Value, PPV)
- Recall (or Sensitivity)
- Specificity
- F1 score
- Root Mean Squared Error (RMSE)
- Coefficient of Determination (R-squared)

When applied in context, developers and implementers can use these considerations, methods, and metrics in conjunction with a series of tools to assure all CHAI principles are applied. A concerted approach to using applied model cards, quality assurance functions, and registries is key to ensuring that high quality AI solutions are used and maintained in a proper manner.

Model Cards: In order to help the community move toward standardization of transparency information shared publicly with customers of AI and aid in streamlining procurement conversations, CHAI utilized the categories found in **ONC's HTI-1 Rule** describing predictive decision support intervention tools and recent **FDA Draft Guidance** to create an open-source, template **applied model card** for public use.

Acceleration of Development: Ensuring an AI solution is trained and validated on high quality data is a key factor in responsible AI. To support robust training of models and the

generalizability of AI solutions, CHAI is standing up a nationwide network of data platforms (sometimes referred to as quality assurance labs) that enable robust training and validation of AI solutions on multi-institutional data sets. Using these platforms will aid in the acceleration of development and objective testing of AI solutions by enabling access to "regulatory-grade", HIPAA-compliant data, and testing that lasts a number of days. Local tuning and robust monitoring are also supported by some of these platforms and will be key components of ensuring effective governance.

Registries: As a core tenant of the CHAI philosophy, CHAI is bringing the developer and implementer communities together to create a registry storing the applied model cards and to bring a greater sense of transparency to the health community. The goal of the registry is to provide a mechanism for vendors to share transparent information with the public that is updated at defined intervals by the developer. Providing this information in a standardized format will enable easy access to information for customers, clinicians, and patients to make informed decisions with.

> Bonus: More information regarding the Coalition for Health AI **found here**. **CHAI summary** provided by Perplexity.

For a deeper dive regarding governance frameworks, Gemini Advanced v1.5 Pro Deep Research provides still **further information is available** here.

3.3.4. International Medical Device Regulators Forum Risk Frameworks

The International Medical Device Regulators Forum (IMDRF) exists to accelerate international medical device regulatory convergence in order to promote an efficient and effective regulatory model for medical devices that is responsive to emerging challenges while protecting and maximizing public health and safety. Established in 2011 as a voluntary group of medical device regulators from around the world, they came together to build on the strong foundational work of the Global Harmonization Task Force on Medical Devices (GHTF) and aims to accelerate international medical device regulatory harmonization and convergence.

One such set of tools that this group has created is a risk framework, which can help evaluators of medical AI assess the criticality impact of a specific tool when used in a clinical environment, against a backdrop of the country regulations that might apply to them. For example, knowing that a solution works with critical patient situations that inform diagnostic choices, will most likely be considered, from a risk perspective, a Class III medical device.

More information on these risk evaluation frameworks in this **paper**.

4. Balancing Fatigue, Timeliness, & Effectiveness

Clinical AI balancing its weights for the end user.
Image generated by Gemini Advanced 2.0 Flash.

As mentioned above, there are many permutations of individual AI-driven applications already available today. Therefore it is probable that any net new initiative produced by one's organization will need to be able to be crafted and deployed in such a manner as to function well alongside AI-driven solutions which were produced by another party (AI developer within one's institution or commercial source). While analysis does benefit from parallelization, the interaction point for the results must be implemented into a unified structure which enables modular and configurable result management for the end user to adjudicate. Without addressing the need to harmonize results from disparate solutions,

clinical team members may find themselves receiving dozens or hundreds of different interaction points on top of the overwhelming number of alerts from monitors or results for each and every patient. This single concept is a strong contributor to a philosophy and strategy of delivering many solutions via a single/standard implementation method which is able to execute on each of the factors outlined within this document. However, harmonizing results from disparate inputs is not the only aspect related to parallelization one should give consideration.

Time is another important factor. For instance, not only does all of the relevant clinical data need to be exposed to one or more models in order to provide the clinician(s) the optimal feedback, those models each must be run at the same time so they may contribute to a patient's study or encounter together. In this manner, the most complete and timely feedback can be delivered to the appropriate caregivers per the use cases involved when that physician needs it in order to be the most effective, improve an existing workflow, and/or benefit a patient. A good example of this would be a triage and detection algorithm designed for detecting an aneurysm of any vessel on a CT Angiogram. Because aneurysms are often asymptomatic, most studies of this type would be ordered for a different indication(s). Any other algorithm deployed and available to address the indication of this study would need to run concurrently to that of an automatically triggered aneurysm algorithm. Many would need to run at the same time on the same single case so that all results (positive or negative) are made available to the user reviewing that CT Angiogram before they get to the study within their workflow. Further, if one were to deploy a clinical algorithm focused on the patient's record in the EMR to complement the finding of an aneurysm, the positive finding from image interrogation would need to serve as the trigger to activate the clinical algorithm (i.e. to summarize relevant parts of the patient's history, ailments, allergies, current labs/vitals, and risk assessment for rupture) and then also return the output of this analysis in time to be consumed with the original imaging based finding in a unified interaction layer. These functions must run in parallel so the user perceives them as a single process.

In image based analysis scenarios another way to think about this phenomenon is in terms of the amount of anatomy on an imaging study which could possibly be mapped to the sum total of pathologies or abnormalities which could be present within that anatomy from that modality. If some day a solution or collection of solutions are available which can detect, measure, alert, etc. for all the relevant pathology, each must contribute independently without interfering with another's function and prior to a physician being ready and capable of interpreting the imaging study in question. In other words, they will all have to be run at the same time. Should clinical algorithms be available, as with the aneurysm example, for each of the total number of possible pathologies detected on imaging, they will all need to be triggered, applied, and synthesized before the end user(s) access the native or raw study/data.

Therefore, any net new solution produced will need to fit into and not inhibit this new paradigm of parallel applications. GPU processing functions perform wonderfully in terms of running multiple analyses at the same time. However, in order for one's net new application to achieve production adoption the activation of that application, the data analysis it provides, and the synthesis of its results also need to be performed in parallel to any other relevant application in a short enough time frame according to the acuity of the pathology involved in order to provide meaningful function to a clinician and/or their team.

As with any rule, there are going to be exceptions to the above section. Today we have examples of single analysis directed to dedicated clinical studies and data types which are widely adopted. One should expect, particularly in the short term from this publication, for the number of these 1:1, dedicated analyses to single AI-driven use cases to grow in a similar manner as today's AI market has grown. Appropriate examples of this which are well adopted today are found in mammography, for instance. A screening mammogram is 'screening' for cancer and its early signs. Therefore, a simple mapping of a mammogram study to a cancer screening algorithm is an example of an appropriate 1:1 match of data to its corresponding algorithm. Another current example is found in the dedicated analysis of coronary artery atherosclerosis. Companies like Heartflow and Cleerly (both originating from the academic community) solely focus their analyses on just a single output for a single disease on a single study.

The recommended focus prior to undertaking an AI-driven project is to identify as early as possible, whether or not one's project will fit into a Parallel or Dedicated category.

If it is best suited for a Parallel workflow, then the challenge becomes primarily a technical one to implement it in a way which is consistent with the examples given in this section above. If it is appropriately designed as a dedicated analysis, then the quality and value it delivers must be commensurate with the time and energy required by the end user(s) to afford the introduction of a novel, net new user interface, report, or content of some kind. As one may already be familiar, the value which must be generated associated with this latter option is quite high.

Finally, once a solution is designed and implemented in such a manner as to be fast enough and fit in with everything else which may need to be simultaneously active, a solution will be judged by its effectiveness. Meaning, does it deliver enough of an improvement over the prior clinical norm which can be trusted? As with the value previously mentioned, the technical demand in order to achieve this level of effectiveness is also quite high.

4.1. Becoming Useful, Usable, and (ultimately) Used

Since the dawn of software (and likely before), vendors have been burdened with the task of getting their clients to fully adopt what they have designed, delivered, and deployed. It's one thing to come up with a novel way to do something which is more efficient than a prior method. It is still another to craft that method in such a manner to make it feasible for many people to access and leverage it. But it is an entirely different endeavor to get other human beings to actually use it in their daily lives and/or at work. This phenomenon in healthcare is called being useful, usable, and then used in a production clinical practice and it is hard to accomplish. In this section, are a few of the requirements in order to attain the coveted status of robust clinical adoption.

4.2. Improve the Clinical Status Quo

As one may have already imagined, introducing a novel piece of technology into a group of clinicians is a task with a high bar for adoption. This is primarily due to the following combination of attributes to the clinical setting:

1. **Lives are on the line:**
   This one is easy to understand. If any solution being introduced could impact acute, life or death scenarios for patients, it presents the potential of both high value and terrible cost. This paradox is an appropriate reason to proceed cautiously when presented with such choices. One's clinical audience is all too aware of this as they live it every day. Anyone developing a tool for clinical use should thoroughly understand what it is like to walk in the shoes (or wear the scrubs!) of the clinical team involved so that the audiences' clinical considerations are fully understood. Whatever is designed must make their clinical life (and vicariously the lives of their patients) better or it will be summarily rejected.

2. **Medicolegal liability can be exquisite:**
   When introducing a new tool to clinicians, one question which one should expect the audience to ask is whether or not the change being proposed increases or decreases their liability. For this reason, as well as the others presented in this section, a prominent filter through which clinicians evaluate novel tools is that of risk mitigation. They will demonstrate a natural aversion to anything which may be perceived as doing anything but lowering their risk exposure. As one will likely see when embarking on the delivery of new tools, change alone is often all that is needed to trigger one's sensitivity to risk aversion. Therefore, the clear understanding and intentional design into a new tool of how this tool, once adopted, would lower clinical risk exposure can be a powerful feature or set of features for one's audience of adopters.

3. "I already do that":

Clinicians of all kinds have dedicated a minimum of years of study, in most cases 10 or more, toward their chosen field, countless hours, hundreds of thousands of dollars in education, years of their life they cannot get back, etc. toward their clinical competency, their expertise. It is not just what they do. It's quite often, what they are. Reasonably so, it's an identity worn as a badge of honor and we are indebted to them for it. However, when this is the audience to which new tools are introduced, hesitation is a common response that should be expected. This is precisely because they have already, by the nature of what they do and how thoroughly they are trained, put people and processes in place designed to account for and address expected and unexpected clinical challenges of all kinds. I recommend one to shadow teams of people in an emergency department if one has never had the opportunity to observe teams of people, working in concert to react expertly to whatever comes through their doors, 24/7, every single day of the year. Therefore, the introduction of a new tool, an adjustment to the sheet music of their concert, must improve the quality of their 'tune' or it will not get a second look. They will simply say (either to themselves or out loud), "I/We already do that." So the goal of an AI-driven tool introduced to clinical settings must be to enable these teams of actual experts to do something they cannot do on their own. The goal is to give the new capabilities, new powers. This will create new value, and that they can justify adopting, using.

4. Margins are quite thin:

Not only is the clinical margin for error very narrow, so are those for the finance teams of each service line into which new tools are presumed to be introduced. This scenario is a natural circumstance which leads to a kind of fiscal and cultural conservatism which is challenging to overcome. Each clinical service line leader who would almost certainly need to review and approve of any new, care-altering tool also has a business administrator related to that service line who is responsible for the economics of that team. If care is to be altered by a tool, economics should be expected to be altered. Healthcare margins are routinely, when healthy, measured in the low single digits of percentages. This means there is very little room for the introduction of anything which alters the method of providing care. ***Therefore, the economic impact of one's AI-driven tool needs to be as well understood as any other facet of how that tool functions and performs.*** It is simply a reality of performing in healthcare today, the clinical/fiscal change made possible by a new AI-driven tool must be an augmentation which is larger than the appropriately protected inertia of that clinical circumstance.

5. Lack of time:

As one may already be aware, we all are currently experiencing a global shortage of clinicians of all kinds. The only variable from country to country is how severe that shortage is among which type of clinician. This very challenging environment is the one in which a novel AI-driven tool must be carefully evaluated by those in leadership, presiding over the service

lines which are understaffed. Those clinicians and IT teammates have the task of weighing all that is outlined in this document in their spare time they don't have. Therefore, careful, thoughtful planning should be made by the innovation team who is designing a novel solution to communicate, gain feedback from, educate, collaborate, etc. into their extraordinarily busy calendars. Time was already the world's most precious commodity. Within healthcare today, it is the highest its premium has ever been.

4.3. Prevent the Avoidable

Timeliness, not to be confused with the extra time just covered in the above, of the delivery of an AI-driven insight will be a significant factor to the perceived value of that insight by its target audience. If a solution is terrific at correctly identifying a meaningful finding on a patient, or even a cohort of patients, but that finding isn't delivered fast enough in the context of that patient's care, the value of that insight can be greatly diminished or even viewed as a complication to appropriate care. In order, therefore, to positively contribute to patient care, to prevent an avoidable negative outcome, one will have to thoroughly understand the optimal delivery mechanism of their AI-derived feedback and ensure how it can be predictably delivered,consumed, and leveraged by its target audience. In addition, if the insight is valuable and users are not consuming that content as intended, one will need to ensure there are the necessary analytics in place to identify how often that is occurring so additional mitigation and/or change management steps can be taken to provide optimal opportunity for the clinical audience to leverage AI-derived insights and avoid the preventable, negative outcomes one's AI solution is designed to lessen.

4.4. Meet them where they are

The 'them' in this case are your end users. As has been outlined, time, expense, clinical demand, etc. all work together to create an environment in which one's target audience for a new solution simply has very little additional attention to give to something which does not already fit directly into how they go about their work each day. Therefore, equal time and/or development emphasis would be appropriately placed on both what it takes to create new AI-insights as well as how and where those insights can be delivered so one's end user audience can fit it as easily as possible into their current workstream. For this to be succinctly accomplished, heavy emphasis will need to be placed on integrations directly to the applications which they are already using. It should be noted that novel or native integrations to production clinical systems can be complex and are almost always an expense that needs to be given careful consideration. The topic of integrations, their role, and some examples will be covered further in the next section of this document.

5. Organizational and Strategic Considerations

5.1 Cost to Produce a Solution

Not quite Thomas Sowell sitting in front of definitely not his students. Image generated by Grok3.

Thomas Sowell, a prominent economist, author, and professor, has asked throughout his career and publications, “At what cost?” We too must be mindful of the costs associated with the development of novel AI solutions for our space. Similar to the regulations applicable, costs can be significant but are known. Because they are known, they can be accommodated. The following data in this section is directional in nature rather than a precise account of what one's costs will be for a given product. The specific costs will be determined by the factors outlined below. Any dollar amounts provided here are simply estimates which will need to be made solution specific.

As you will see, the cost estimates for each of these sections can vary widely. We regret to relay that the rather unsatisfying answer to the reason for these variations is, it depends. Model size, modality involved, prevalence of the pathology involved, goals regarding reproducibility, regulatory involvement, and the varieties of skilled labor needed to execute on each of these needs can easily create a seven figure swing in the budget for a project. For instance, if one is designing a model to detect a single, prevalent pathology on an X-ray

modality, the costs incurred could be on the lower end of what is estimated. However, if one prefers to target a lower prevalence disease which is only found on MRI case, it is likely this could be (due to lower disease prevalence, probability of an existing 510K clearance, and higher de novo costs outlined in section 5.1.5.) one of the more expensive models to try and create. The latter's needs on data curation, computation, storage, transfer, annotation expertise, etc. will all be much more than the former.

The same can be said for language based AI models. The lower the prevalence of what is sought by the model, the larger the data set it would need to be trained on in order to achieve the output necessary to be trusted by its audience. For clinical language based models one has the added sensitivity of needed access to large swaths of EMR data on which to train and be validated.

5.1.2. Data Collection: $300,000 - $5,000,000

It may sound straightforward, to collect the data needed for analysis. However, the devil is in the details on this issue. This isn't just 'data' that can be gained from a variety of sources. The subject here is HIPAA protected, expertly annotated, securely transferred and stored patient identifiable healthcare data. Each of the attributes just mentioned require their own layer of expertise in order to handle properly. In addition these 'requirements' may vary by jurisdiction. Meaning, Canada may have a different set of security and privacy requirements as the EU vs. the United States vs. India and so on. If one is planning on creating a model of any kind which applies well to the diverse populations just mentioned, collecting enough data in a secure and compliant manner from those locations will require a full understanding of the needs particular to that part of the world.

Another meaningful variable for which to account is the size of the data set in question. Size will determine things like storage needs or the time it takes to execute the transfer, both of which have cost implications. Size may also be an indication of the amount of data which needs annotation - either on imaging or clinical records. Annotation, the process by which labels or metadata are added to the raw data such that it can be understood by the algorithm being trained, will need to be completed by experts from the applicable clinical domain. This is a scarce kind of labor and will come at a premium in most cases. However, performing it well is critical to the success of any AI solution.

5.1.3. Algorithm Development, Evaluation, and Testing: $200,000 - $3,000,000

The development of complex deep learning models based on convolutional neural networks (CNN), can require significant computational resources and specialized expertise to perform. It is not unreasonable to need a solution architect, whose task it would be to oversee and balance both the business and technical aspects of model creation, and often have an income expectation of $250,000/yr (7).

In order to then ensure the resulting algorithm meets the performance and acceptance criteria warranted by a production, clinical environment, one should expect to need to run extensive experiments in order to refine the model to this level of reliability. This will not only add to the time involved in the overall project but also contribute to the computational demands already mentioned. One strategy worth considering to mitigate some of this workload is to begin one's project with prebuilt tools, such as that of the MONAI open source community. Their open source toolkit is **available here**. However, if one's project requires a customized start, expect the development costs to raise significantly to the higher end of the quoted range above.

5.1.4. System Design: $250,000 - $4,000,000

As has been covered throughout this document, the system needed to run clinical AI solutions well and in production is significant. In addition to ensuring proper activation and monitoring performance for drift mitigation (which one can think of as necessary ends of a heavy barbell of solutions), a sufficient user interface (UI) will need to be found (if an existing one can be leveraged) or created. There are entire subspecialties of application development dedicated to elegant user interface design. One should expect the degree of emphasis placed on this design to be directly related to the costs associated with producing it. While there are a plethora of automated application coding solutions available today, one would still need to acquire the skillset needed to perform this task well. A difficult to use or poorly implemented UI is one of the fastest means of limiting the impact of a piece of technology. However, one may consider money invested in elegant design to be well spent as that design is correlated to a well adopted solution, which is one of the main themes of this entire document.

In addition to the above, the workflow required by a solution will also need serious deliberation. One can easily find that the need to deliver the best AI output to the correct clinician at just the right time on the correct patient encounter or event and within a native and/or elegant new UI to be an exquisite challenge. Therefore, extensive discovery and understanding of one's targeted users' needs as well as the needs of their patients will pay

dividends in avoidance of a lengthy iteration process in solution design. It should be noted, however, that even if well understood, the task of creating the workflow needed may still be extensive, requiring a team of experienced and skilled professionals.

5.1.5. Regulatory Clearance: $500,000 - $1,800,000

As was covered in Governance section 4 of BRIDGE, the regulatory requirements for a clinical AI solution can be significant. Below you can see how some of those costs can come in several forms. Depending on the goals of one's solution, one may have the need to execute each form of regulatory approval.

Consider the following scenario. An academic medical center (AMC) establishes an IRB to create a de novo solution to detect a rare disease on brain MRIs. This solution is designed, tested, and proven to work. In fact, it works so well, the AMC decides they are going to acquire and train this solution on a larger scale so they can pursue FDA clearance. This clearance is granted which allows them to establish a company to deliver this tool to other organizations outside of the AMC within the United States. Because they did such a fantastic job, their solution proves to be reproducible at other organizations and benefits many thousands of additional patients. This success leads to the eventual need to proceed with approval from the European Commission (CE marking). While the FDA and EU present the largest and most immediate paths to business growth, they do not represent regulatory costs in their entirety as Canada, Indonesia, India and so on all have their own regulatory infrastructures which eventually will require navigation for our unnamed but rather successful small business which is growing to a larger one as a result of their successful solution design. But, because of this successful path, they would have to pay each of the regulatory fees listed below and highlighted in green. This can obviously add up.

Annual Establishment Registration Fee:
$9,280 (This fee is paid by all establishments, and there is no small business reduction.)

510(k) Premarket Notification:
Standard Fee: $24,335
Small Business Fee: $6,084

De Novo Classification Request:
Standard Fee: **$162,235**
Small Business Fee: **$40,559**

To reach clearance for being CE Marked, the path is known but less clear (pun intended). The unnamed company above will have to enter the following process and understand the ranges which can be found with these costs:

Within the European Market one will need to comply with the EU Medical Device Regulation (MDR). This requires one to:

Select a Notified Body associated with the classification of the device involved. The more complex the use case, the more the costs are elevated with the complexity of the algorithm, the validation process, and scope of the necessary review. Accordingly, it is all but the lowest classification, Class I, which requires Notified Bodies - the Class IIa, IIb, and III SaMD.

Depending on the device classification, the associated certification process will include three distinct parts: a clinical review, a technical file review (which would include any software validation) and a quality management system audit. The fees incurred in association with this process go beyond just the submission but also the ongoing compliance with validation, audits, and ongoing surveillance of the solution which creates ongoing costs. Should any design changes be required in reaction to this review process, this will incur additional costs. Due to the complexity of this process, it is reasonable to consider engaging with a consultant who specializes in this domain. While helpful, this will also add to the costs.

Thus, for the high risk solution of our unnamed company, becoming and maintaining CE Marked status can range from **€150,000 to €500,000** just in the regulatory related fees. This brings the regulation fees alone to +$700,000 on the upper end. They could go significantly higher, depending on the actual complexities of a real world project which can involve complex clinical study design, protracted review periods, and any fee changes from Notified Bodies (they are subject to changes).

5.1.6. Post Market Surveillance $250,000 - $2,000,000

Not only is postmarket surveillance necessary when going through a regulatory approval process, “surveillance” has gone by the term “drift mitigation” elsewhere within this paper. As has been covered, this is a critical component of the minimum viable product environment needed for an AI solution to be used well over time. Therefore, the solution design process which is directed to a project that would require regulatory approval will require an analytic solution which could be used in the regulatory approval process as well as the ongoing maintenance of the solution for its use in production.

The primary function of this surveillance/mitigation service is to maintain the expected output for the solution in question. However, a secondary byproduct from having the ability

to analyze how well an AI solution is functioning is to also leverage that information in the assessment of its value to the system within which it is deployed, the clinicians who use it, as well as the patients affected by its function. This is another critical component of the successful deployment of an AI tool. In today's healthcare environment, the need to justify the cost of any net new project is ever present. Therefore, designing at the outset of a new project precisely how the value of that output of that project will be viewed will pay dividends once a successful project is ready to be brought live in a production environment.

It may be expected that, in order to deploy such a solution or process which can function at an enterprise scale, that one will need to deploy a combination of technologies in order to execute well. Commercial or open source Natural Language Processing (NLP, another form of AI) tools could be used to identify the instances on which an AI solution took action as well as compare that action to its corresponding ground truth in the clinical record. However, as was covered previously in this document, prevalence will play a strong role as to whether or not NLP tools are accurate enough to be solely relied upon for assessing an AI Solutions function. If automated tools like NLP cannot be statistically relied upon, other means or proxies to expected function will need to be identified and tracked accordingly. Once the means of tracking the data have been determined, easy visualization of that data will be needed. Commonly used dashboards from companies like Tableau are often used for this purpose. It is probable that a health system may already have access to such a tool (or a similar one) which would be satisfactory for the needs outlined here.

5.1.7. Preparing to Scale $300,000 - $5,000,000

If one has satisfied the above in this section, congratulations may be in order. You have accomplished something which would make Jimmy Dugan smile. However, the next task is also significant. That is, taking your existing project, at whatever scale it currently resides, and expanding it to an entire enterprise. Many health systems today have tens or dozens of facilities. They are large and they employ thousands or even tens of thousands of clinicians. How many of them does your solution need to reach if it were to be completely adopted? This will need to be very well understood and accommodated, which is reflected in the infrastructure powering and running the workflow for your given solution. The planning and execution of this will involve but not be limited to:

- Network capacity
- VM/server specifications
- Enterprise system integrations
- Cloud Services
- Model quality
- System analytics
- Accompanying test environment
- Security measures
- Regulatory compliance

In taking the whole of what is outlined in this section together, one should expect to need a sizable investment in order to implement just a single solution: minimum $1.8 million, taken from the lowest end of the estimations. If one's goals are more extensive, perhaps targeting multiple, reproducible solutions, the minimum costs could exceed $20 million.

More information, regarding cost and benefit analysis for clinical AI tools are provided by Gemini Advanced v1.5 Pro Deep Research and **can be found here**.

5.2 Reimbursement Landscape

5.2.1. The Reimbursement Triangle: Codes, Coverage, and Payment

Reimbursement for artificial intelligence (AI) in clinical practice is a complex and evolving landscape. It hinges on three interdependent factors: codes, coverage, and payment.

1. Codes – The foundation of reimbursement, codes classify medical services and procedures. These include Current Procedural Terminology (CPT) codes, Diagnosis-Related Groups (DRGs) under the Inpatient Prospective Payment System (IPPS), and Ambulatory Payment Classifications (APCs) under the Outpatient Prospective Payment System (OPPS).

2. Coverage – Determines whether payers (e.g., Medicare, Medicaid, private insurers) will reimburse for a given service. AI solutions must demonstrate clinical utility and cost-effectiveness to be included in payer policies.

3. Payment – Defines how much providers are reimbursed for AI-enabled services. Payments can be bundled (e.g., DRGs for inpatient care) or provided through add-on payments such as the New Technology Add-on Payment (NTAP) or New Technology APCs for emerging innovations.

AI solutions often struggle to fit neatly into existing reimbursement models, requiring strategic planning to secure both regulatory approval and financial viability. This section explores key U.S. reimbursement pathways and how recent policy changes are enabling AI technologies to integrate more seamlessly into the healthcare system.

5.2.2. U.S.: Key Pathways - Inpatient Prospective Payment System (IPPS) and NTAP

The IPPS, managed by the Centers for Medicare & Medicaid Services (CMS), reimburses hospitals for inpatient services based on Diagnosis-Related Groups (DRGs). AI technologies used in inpatient settings—such as decision support for stroke triage or AI-powered imaging tools—are typically included in existing DRG payments. Since DRGs are bundled payments, hospitals must absorb the cost of AI unless it qualifies for additional reimbursement.

One such mechanism is the New Technology Add-on Payment (NTAP), which provides temporary supplemental payments for new medical technologies that are not yet fully incorporated into DRG rates. To qualify, a technology must be:

- New (typically within three years of FDA clearance),
- Above a cost threshold, meaning its use would create a significant financial burden for hospitals under existing DRG payments,
- Clinically beneficial, demonstrating substantial improvement over existing treatment options.

5.2.3. Outpatient Prospective Payment System (OPPS) and New Technology APCs

For outpatient settings, OPPS governs reimbursement, assigning procedures and services to Ambulatory Payment Classifications (APCs). AI applications used in imaging, diagnostics, and procedural guidance typically fall under existing APCs.

For AI-driven services that do not yet fit established APCs, the New Technology APC pathway offers a temporary solution. AI tools demonstrating significant clinical impact—such as AI-based cancer detection algorithms—can receive an APC-specific payment until a long-term coding solution is established.

However, while New Technology APCs provide a temporary bridge to reimbursement, they do not establish long-term financial sustainability for AI applications. To secure permanent reimbursement, AI developers must navigate the **CPT coding system**, which dictates how services are classified and billed across healthcare settings.

5.2.4. Current Procedural Terminology (CPT) Codes: CPT I vs. CPT III

CPT codes define reimbursement eligibility for medical procedures and services. AI solutions seeking direct reimbursement must align with the CPT framework:

- **CPT I Codes** – Reserved for well-established procedures with widespread clinical adoption. AI-based services rarely obtain CPT I status initially, as they must demonstrate FDA clearance, broad clinical use, and payer support.
- **CPT III Codes** – Temporary tracking codes for emerging technologies, including AI-powered diagnostics and risk assessment tools. These codes allow data collection but do not guarantee reimbursement, as most payers wait for clinical validation before covering them.

CPT III codes are often a stepping stone toward CPT I classification, requiring developers to generate real-world evidence that proves both clinical efficacy and cost-effectiveness.

5.2.5. Integrating Reimbursement into Early Product Life Cycles

Historically, reimbursement has been an afterthought in AI development, addressed only after regulatory approval. However, new programs—such as the FDA Total Product Life Cycle Advisory Program (TAP)—enable AI companies to discuss reimbursement earlier in the process, reducing delays in market adoption.

5.2.6. FDA Total Product Life Cycle Advisory Program (TAP)

The TAP program provides structured engagement between medical device developers, the FDA, payers (e.g., CMS, private insurers), and healthcare providers. AI companies can:

- Identify reimbursement barriers early, ensuring AI adoption is financially viable.
- Align clinical trial endpoints with regulatory and payer requirements, streamlining coverage decisions.
- Optimize study designs to collect evidence needed for both FDA approval and reimbursement.

By integrating payer perspectives early, TAP increases the likelihood of AI technologies securing both FDA clearance and sustainable reimbursement pathways.

5.2.7. Breakthrough Device Designation and Its Impact on Reimbursement

The Breakthrough Device Designation (BDD) program accelerates FDA review for devices that address life-threatening or irreversibly debilitating conditions. AI solutions for stroke triage, acute pulmonary embolism, or cancer detection often qualify.

Key benefits include:

- Priority FDA review, reducing time to market.
- Streamlined clinical trial requirements, lowering evidence burdens.
- Favorable NTAP consideration, as Breakthrough-designated AI tools face reduced requirements for demonstrating clinical benefit.

Breakthrough designation can also influence Transitional Coverage for Emerging Technologies (TCET) policies, potentially enabling immediate Medicare coverage upon FDA approval.

5.2.8 Reimbursement for AI Outside the U.S.

International markets present challenges specific to their markets:

- **Europe (EU)** – Reimbursement varies across countries, with Diagnosis-Related Groups (DRGs) used in some nations (e.g., Germany), while others rely on individual payer negotiations. AI adoption is often hindered by fragmented regulations and slow CPT-equivalent coding updates.
- **United Kingdom (NHS)** – AI reimbursement depends on National Institute for Health and Care Excellence (NICE) evaluations, requiring strong evidence of clinical benefit and cost-effectiveness before widespread adoption.
- **Japan** – The government actively invests in AI-driven healthcare, but reimbursement depends on approval from the Central Social Insurance Medical Council (Chuikyo), often requiring multiple years of negotiation.
- **Australia & Canada** – AI reimbursement follows Medicare-type models, where government agencies dictate coverage, leading to lengthy review cycles and limited AI-specific payment pathways.

Global AI adoption requires a targeted market approach. While globally fragmented, proactive AI reimbursement strategies can accelerate market adoption. By addressing codes, coverage, and payment early—through programs like TAP and BDD—AI developers can reduce regulatory and financial uncertainty. Temporary mechanisms like NTAP and New

Tech APC provide interim reimbursement, long-term success depends on achieving CPT I status and integration into standard payer policies. As AI's role in healthcare grows, aligning clinical validation with reimbursement strategy will be essential for ensuring that innovative solutions become both clinically impactful and financially sustainable.

5.3 The Potential Impact of AI on Medico-Legal Liability

As with any new technology or clinical pathway adoption the use of artificial intelligence in healthcare introduces new complexity for implementers. One such complexity to consider is that of medico-legal liability. Liability is directly related to the likelihood in a clinical scenario a patient or their family would pursue a lawsuit, following a care encounter. In this section we'll explore those motivations as well as point to what we believe are some strategies to further mitigate risk.

There have been many papers and abstracts published, showing the positive impact clinical AI solutions can have on patients. Wiklund, et al. showed a dramatic improvement in pulmonary embolism detection on oncology patients and Armoundas, et al. outlined a number of use cases with positive impact for heart failure, sepsis, ECG, and perioperative risk assessments (8), (9). Cedars Sinai authors highlighted downstream effects, reducing hospital length of stay for patients who had triage solutions applied to their care (10).

However, the entry of AI into the delivery of healthcare might also increase end-user (i.e., healthcare providers' and health system's) risk related to medico-legal liability in specific circumstances. For instance, there are known cognitive biases, such as the "automation bias," wherein the subject's level of trust in the machine's diagnostic accuracy is essentially unbounded, that can influence the outcome of litigation in tort cases alleging medical malpractice where AI has played a role in clinical decision-making. In the foreseeable future one can readily envision scenarios where a clinician's response to AI guidance could become the sole basis of a legal claim. This could call into question what the applicable care standard is, or should be, for clinicians who utilize AI tools. There are, however, a few hopeful signs that these new risks can be effectively mitigated.

For perspective on the magnitude of this problem, we note that nearly half of physicians 55 years or older have faced at least one medical malpractice lawsuit claim in their career (11). The average physician spends more than four years with an open malpractice claim (12). Artificial Intelligence (AI) will change the legal landscape for providers in a way that is both presently unrealized and evolving (13). AI could modify liability by impacting one or more of the following in the event of a bad medical outcome:

1. A patient's desire to pursue legal action
2. A lawyer's willingness to accept a medical malpractice case
3. If a case is settled: what the settlement amount would be
4. If a case goes to trial: what the judge or jury verdict would be

Recent research from the **Brown University Radiology Human Factors Consortium** suggests that, for items 1 and 4, there is indeed risk for radiologists who use AI tools. However, the reality is more nuanced than it may seem. When it comes to false negatives, the effect of AI on legal liability would likely depend on whether:

1. AI also made an error of omission; and interestingly
2. Whether the AI error rate was known

In cases when a radiologist failed to detect an abnormality that resulted in eventual breast cancer on a screening mammogram, women aged 40+ reported a moderate to large increase in the likelihood that they would consult with an attorney to consider suing the radiologist if AI correctly flagged their case relative to if no AI were used (14). Similarly, when mock jurors were asked how they would decide a case where a radiologist was being sued for a failure to detect an abnormality, they were more likely to side with the plaintiff if told an AI system was used which correctly flagged the case (relative to no AI being used). However, when AI also failed to detect an abnormality, the AI penalty largely dissipated (15).

Fortunately for physicians, current research also demonstrated a protective effect of making people aware of the AI's actual error rates, showing that revealing this information to the patients and mock-jurors had a strong impact on their conclusions regarding the physician's legal liability. In the former study, when AI correctly flagged a case that a physician missed, women reported a much lower likelihood of consulting a lawyer for a medical malpractice lawsuit when told (versus not told) that 95% of the cases flagged by AI are actually false positives. In some cases, the legal liability was even lower than when no AI was used. Put differently, a physician's liability for incorrectly disagreeing with AI is eliminated when providing a patient with the false discovery rate of that AI.

Likewise, when mock jurors in the latter study were provided with AI error rates, the percentage who sided with the plaintiff for a missed brain bleed decreased drastically. This occurred for both providing the AI false discovery (FDR) and false omission rate (FOR) when AI missed and caught the brain bleed, respectively. Examples of this sort demonstrate the importance of transparent information sharing between solution providers, implementers, and patients. Model card and "nutrition label" type initiatives have already gained steam with organizations like The Coalition for Health AI (CHAI) launching open-source applied model cards, Sendak, et al having published "model facts" labels, the FDA offering "Transparency Design Considerations" encouraging the use of model cards in their January 2025 draft

guidance, and many others. While many model cards will be published publicly, an emerging frontier is a patient facing summary. Similar to the Journal of the American College of Radiology (JACR) patient-friendly summaries of appropriateness criteria guidelines, these summaries could be important tools for patients and caregivers to understand the real-world usage of these cutting-edge solutions, provide an understanding, and establish trust.

In summary, although preliminary, the available data collected to date suggest:

- AI may increase legal liability, especially when they contradict or disagree with/reject the AI guidance, and
- This effect is strongly mitigated, or even eliminated, when people are provided with data on the frequency with which AI makes mistakes, thereby controverting any "automation bias" that may have otherwise swayed these subjects.

It seems prudent, therefore, for clinicians to consider always reporting the FOR and FDR of any AI that is being used, as a matter of policy, which would require the associated analytics to be readily available.

Finally, we must point out an unknown factor at this time. That is, as more well adopted solutions are deployed, one plausible outcome may be that healthcare in total experiences a decrease in its overall error rate. For instance, Rad Partners, the largest independent radiology practice in the United States, has demonstrated enterprise wide enhanced detection rates for intracranial hemorrhage (12.6%), pulmonary embolism (18.1%), incidental PE (35.8%), and fractures of the cervical spine (16.4%) or rib (60.5%) (16). Were this kind of effect, as well as those previously mentioned from Armoundas, et al, to be broadly experienced across the vertical, we may observe a reduced total pool of missed pathologies which could lead to a lower volume of the kind of medical mistakes outlined by Newman-Toker, et al. in which the research detailed how 38.7% of most common medical errors which lead to patient harm are related to just five pathologies: stroke, sepsis, pneumonia, venous thromboembolism, and lung cancer (17). Were a concentration of well adopted and performing AI solutions to target those pathologies, a meaningful change in the medical error rate may be measureable, which may relate directly to the volume of instances of medico-legal liability.

6. Regulatory and Compliance

"Governance structures to withstand time." Image generated by Gemini Advanced 2.0 Flash.

With BRIDGE, we are attempting to summarize the regulatory considerations of an AI model, its developer(s), and the process in which it is developed need to be aware and/or take account so the AI model(s) in question can comply with regulations applicable to their jurisdiction over time. BRIDGE endeavors to create alignment on terminology and definitions as well as a summary for each of these regulations. These include TEFCA, IRB, ACR, ISO, NIST, and others (see section immediately below). To our knowledge, BRIDGE makes the first attempt to collate and summarize all those applicable regulations within a single document. Following each section, one will note that we have provided reference links to the actual regulations in question.

The following discussion is neither designed nor intended to be a substitute for professional compliance or legal advice. All are encouraged to speak with their counsel and are exclusively responsible for ensuring solutions comply with applicable Federal/State laws and regulations.

6.1. HIPAA - Health Insurance Portability and Accountability Act

The Health Insurance Portability and Accountability Act is intended to both protect patients' health information as well as provide them more control over how it is used. The Act became Federal Law in the United States in 1996, and it established the standards for appropriate utilization of healthcare data.

Due to its age, HIPAA has a well worn path of legal and safety regulations being implemented in the standard practices of today's US based health systems. Therefore, one will need to understand those existing policies before exploring any new paths needed to both develop new tools and remain compliant. For example, your organization likely will already have a Data Governance Policy which references HIPAA. Therefore, it is very likely the means of compliance with this twenty-eight year old US Federal Law (or GDPR - General Data Protection Regulation - its equivalent in the European Union) is to simply follow established policy.

Ensuring this compliance will help clear an early hurdle in solution design as well as provide alignment with one's internal policies. This is a procedure which will repeat itself as one goes through the process of ideation, design, training, testing, iteration, and (hopefully) implementation. Along the way, there will, no doubt, be what seem to be endless committees through which one's solution must be assessed and reassessed against HIPAA and the other regulations outlined within this section.

Additional detail regarding the Healthcare Insurance Portability and Accountability Act **found here**.

Also, this is a **HIPAA summary** provided by Perplixity.

6.2. IRB - Institutional Review Board

For solutions designed and used within your own organization only, an IRB (Institutional Review Board) may be required. In general, IRBs play a vital role in ensuring that any AI solution deployed for clinical research would adhere to one's organizational policies related to patient ethics, safety, and with the necessary institutional oversight . IRBs have several key components, involving the nature of its members, the purpose of the board, as well as mechanisms for continual improvement.

Example Review Process:

The IRB will often establish a process for reviewing research protocols to ensure they meet ethical standards and regulatory requirements. This typically involves:

- Initial review and assessment of risk
- Informed consent to ensure protection of the rights of subjects involved
- Convening meetings to discuss and deliberate on protocols which warrant it
- Providing researchers with feedback and requesting modifications
    - Approving, disapproving, or requiring modifications to protocols
- Ongoing monitoring of approved research
- External IRB options also exist for systems that may not have an IRB structure already in place
    - Western Institutional Review Board-Copernicus Group (WCG® IRB) is an example of an external IRB that has achieved ISO 9001:2008 Certification

> Additional information regarding Institutional Review Boards **found here**. **IRB summary** provided by Perplexity.

6.3. FDA Regulatory Framework and Classification

Should one wish to leverage an AI solution outside of one's 'four walls,' understanding the FDA's classification system for AI medical devices is crucial for both developers and healthcare organizations. This framework determines the FDA regulatory pathway from development requirements through post-market surveillance. The classification system reflects the FDA's risk-based approach to regulation where higher-risk devices face more stringent controls. For developers, early understanding of likely classification helps shape development strategy and resource allocation. For healthcare organizations, classification is indicative of the level of validation and oversight required for safe implementation as well as budgeting and resource demand needed for higher vs. lower classifications.

The FDA currently regulates three types of AI for commercial use in clinical healthcare settings, all classified as Software as a Medical Device (SaMD), which is defined as software intended to be used for one or more medical purposes that perform these purposes without being part of a hardware medical device (18):

1. Image-based AI models that analyze medical imaging studies
2. Clinical Decision Support (CDS) algorithms that process alpha-numeric data
3. Signal Processing Devices (SPD) that analyze temporal/time-series physiological data (waveforms/biosignals)
4. In-vitro Diagnostic (IVD) medical device, capable of running on (non-medical purpose) computing platforms or mobile apps for the purposes of screening, diagnostic or monitoring.

6.3.1. FDA Evaluation Framework

The FDA evaluates all medical AI tools through several key categories for regulatory oversight. These evaluations begin with determining regulatory applicability and continue through intended use and scope assessment. The level of risk classification forms a central component of evaluation along with the degree of automation - whether fully autonomous or requiring human oversight. The FDA carefully considers both clinical impact and data processing methodology in their evaluation process.

Progressive Device Classification and Requirements

Class I Devices - Minimal Risk
Class I devices present minimal risk in healthcare settings and are not used for life support functions. These devices carry no unreasonable illness or injury risk, typically including medical calculators, visualization tools, and non-diagnostic image management systems.

Class II Devices - Moderate Risk
Class II devices present moderate risk and complexity, influencing but not determining clinical decisions. These include CADe (detection) software and monitoring algorithms, encompassing both CADt (triage) and CADe applications.

Change management for Class II devices requires SaMD Pre-Specifications (SPS) and Algorithm Change Protocol (ACP), establishing clear performance boundaries and modification procedures.

Class III Devices - Significant Risk
Class III devices support/sustain life, represent novel technologies without predicates, and/or include diagnostic software for life-threatening conditions. These devices encompass the vast minority of overall solutions which have been designed to date but do include diagnostic (CADx) applications.

Quality Management System

Quality Management Systems (QMS) is the set of policies, procedures, and processes by which an organization maintains quality. Every regulatory authority, including the FDA, places significant emphasis on QMS as they demonstrate systematic control over device safety and effectiveness. Accordingly a robust quality system is not just a regulatory requirement but a framework for maintaining consistent performance and managing change. This aspect of one's solution design is best established before submission and maintained throughout the product lifecycle, making it a critical consideration for both development and implementation strategies.

Performance monitoring systems may include components such as the following:

1. Comprehensive drift detection mechanism
2. Statistical process control measures
3. Root cause analysis protocols, with clear corrective action procedures
4. Incident response protocols defined alongside adverse event management systems
5. Change management procedures documentation and communication methods

While each of those is a significant technical and implementation challenge, one should be encouraged to know that others have been able to implement such systems. Additionally, one should consider what would occur in the absence of such sustainability focused systems and policies: Algorithmic Drift. Drift can be defined as variations in data which, when presented often enough over time, will cause your AI solution to no longer work as designed or initially deployed. As will be covered in a section dedicated to Drift (Section 3.1.4. Mitigation), designing the appropriate QMS system to accompany an AI solution will be paramount to the success of that solution.

Real World Considerations

FDA clearance of medical AI tools confirms basic safety and performance standards within specific validation environments, but doesn't guarantee effective real-world performance, especially when deployment settings differ from training environments.

A potential gap between FDA clearance and real-world effectiveness centers on data representation and reproducibility. Critical factors include training data diversity and its alignment with intended deployment environments, users, and patient populations. The FDA's requirement for clearly stated intended use offers a clear guide for one to understand expected interaction protocols.

Therefore, AI tool development must balance FDA clearance requirements with real-world viability through comprehensive, representative training and validation data. Success requires both regulatory compliance and practical effectiveness, achieved through careful attention to data diversity and environmental factors but will also include both high and low quality study/data examples so as to best represent the irregularities present in real-world clinical environments.

Additional information regarding the Food and Drug Administration **found here**. Also, this is a **FDA summary** provided by Perplexity.

6.4. HTI-1 - Health Data, Technology, and Interoperability

The Health Technology Information Update to the 2016 21st Century Cures act took the form of the HTI-1 Final Rule which was released in January 2024 by the Office of the National Coordinator. Although HTI-1 overlaps significantly with the FDA's regulatory purview over Software as a Medical Device, one of the primary effects of HTI-1 is to create a regulatory framework for those AI models meant for healthcare Certified Health IT vendors or health systems using certified health IT, which fall outside of the FDA's regulatory scope. Most LLMs (Large Language Models), and generative AI - for which there will be endless numbers of use cases and healthcare applications - do not meet the requirements for CDS as outlined by the FDA, but do fall under the HTI-1 definition of Predictive Decision Support Intervention (PDSI). Today, examples of these are decision support tools, administrative or workflow optimization, and natural language processing tools. This is important to note for those who intend to create a clinical solution leveraging a LLM or VLM (Vision enabled Language Model).

In addition, HTI-1 does not have a clearance process like the FDA, but rather requires that systems using a PDSI self-attest that they are doing so. A PDSI is a technology which derives relationships between training data and produces outputs, resulting in predictions, classifications, recommendations, and/or analysis. Attestations serve as declarations that a provider or user of AI is in fact using a specific PDSI tool and understand the requirements for transparency. The attestation also implies that the user of a specific PDSI tool understands the tools intended use.

It provides a definition of AI PDSI as a subset of Decision Support Intervention (DSI) that uses statistical or machine learning methods, and is a broader category designed to help providers make clinical decisions. The typical application of this produces probability scores or risk assessments. Separate from a simple DSI, Predictive DSIs learn dynamically and create outputs that are probabilistic in nature and that depend on large amounts of data and relationships between them. Regular DSIs are much more deterministic and easy to understand how they generate given outputs. With traditional DSIs, it is easier to determine which data is linked to a certain output, and decision pathways are straightforwardly traceable. Predictive DSIs are inherently less deterministic, understandable, and traceable which is why HTI-1 is primarily focused on creating transparency in PDSI training methods, validation methods and results, data sources and source variety on the basis of demographic characteristics, bias mitigation methods, funding source, and more. HTI-1 also outlines the need for a select group of individuals at the health system which is using the PDSI tool to be enabled to change the weights and characteristics of the model to their preferences.

It is, therefore, important to note that the requirements outlined by HTI-1 for PDSIs apply explicitly and exclusively to PDSI's deployed by and purchased through Certified Health IT vendors (a definition related to meaningful use). In most cases, the certified health IT vendor is likely to be their respective EMR vendor. While commercial and "self-made" or "in house" PDSI tools are not legally subject to HTI-1 regulatory framework, many health systems, particularly Academic Medical Centers are requiring that any AI tool which is deployed in their system is compliant with the HTI-1 requirements for transparency as a high-level guide to their internal governance processes. Thus, if your solution or expected project is of the LLM, VLM, or Generative AI variety, close attention should likely be given to the Attestation requirements for use of such a solution in a production environment.

It is also important to outline two additional aspects of HTI-1 which are related to the above and need to be understood. First, HTI-1 outlines what constitutes transparency for Fairness, Appropriateness, Validity, Effectiveness, and Safety (FAVES). With respect to how FAVES are related to PDSIs, they need to account for the following:

Baseline Algorithm Information - developers must provide a consistent set of information about the algorithms used in a PDSI

Source Attribute Categories - these include PDSI intervention details, development insights, and fairness assessment processes

Feedback Mechanisms - modules certified to PDSI need to enable users to select, activate, and provide feedback on both evidenced based and PDSIs.

Second, as of this writing, it is not yet clear the extent to which changes made to the model weights within a LLM or VLM based solution would change the model over all. This is of particular note given HTI-1's requirements, regarding transparency. This will very likely, for the time being and/or until changes in model weight are better understood, mean that straight forward and deterministic use cases should be given primary consideration. What one will want to be mindful of is producing a use case which requires broad application of LLM/VLM technology that cannot be explained. This could end up putting one in the unenviable scenario of producing a solution which seems to provide benefit but is unable to be deployed in a compliant manner because it isn't sufficiently explainable in terms of mitigation, bias, drift, and/or outputs.

> Additional information regarding the HTI-1 Act **found here**.
> Also, this is a **HTI-1 summary** provided by Perplexity.

6.5. ISO - International Standards Organization

While originally voluntary and industry-consensus based, ISO is an independent conglomerate of 167 different organizations, which have worked together to provide requirements and specifications for materials, processes, and products to help make sure they fit their intended application and/or use. Today for SaMD, ISO standards play a critical role in supporting quality management, risk management, information security, and AI governance. These standards help ensure that software solutions used in healthcare meet global regulatory requirements (FDA & CE Marked) and build trust among users and patients. Because the ISO encompasses a broad set of applications, it has proven to be beneficial as a standards body. Many disparate organizations can use these common standards to build and deliver products global customers can trust.

ISO 27001 provides a framework for organizations to establish, implement, and maintain an information security management system (ISMS).As more and more production systems in healthcare, which now even include the EMR, move to the cloud, ISO 27017, which addresses challenges around data location, cloud-specific security controls, and data sovereignty, has become increasingly important. In a related sense, so has ISO 27018, which provides guidance for handling personal information.

With respect to AI solutions, a concert of ISO standards are applicable. While not intended to be a comprehensive list, ISO 13485, 23894, 27001, and 42001 are important as they cover the principles, risk management framework, data quality management, and the deployment, respectively, of AI related systems.

There are two immediate implications of these standards. First, if one wishes to develop a solution for internal use (perhaps under IRB approval), then adherence to the ISO policies already in place. If an existing IRB cannot be used, a novel one will need to be established. However, if a solution will be used outside of one's organization, then the ISO policies in place at all other potential deployments are also relevant. Thankfully, ISO is a well adopted framework across healthcare. Therefore, one should expect consistent experiences once the relevant ISO standards are met for a solution.

AI is rapidly challenging existing standards. One can expect either old ISO standards to be revised and/or new standards which need to be developed/adopted as technology progresses.

> Additional information regarding the International Organization for Standardization **found here**.
> Also, this is a **ISO summary** provided by Perplexity.

6.6. NIST - National Institute of Standards and Technology

The United States National Institute of Standards and Technology (NIST) provides a framework that serves as the practical foundation for how organizations implement standards while meeting U.S. regulatory requirements.

NIST takes a practical approach to defining and managing AI. Their view of artificial intelligence is straightforward - if a computer system can handle tasks that typically need human intelligence, that's AI. This includes systems that can learn from their experiences and adjust their behavior based on new information, much like humans do.

In an effort to help organizations handle AI responsibly, they've created a comprehensive AI Risk Management Framework that serves as a practical guide for developing or using AI systems (19). One can think of it as a roadmap that helps organizations navigate the complex landscape of AI development and deployment safely.

More recently, NIST has been focused on generative AI and foundation models. They have released several guidelines addressing these technologies (20). Their guidance is particularly important because these types of AI can have wide-ranging impacts across different sectors and applications in a domain which is also moving very quickly.

The value of NIST's approach is that it's rooted in real-world application. As a federal agency within the Department of Commerce, they understand both the technical and practical business aspects of implementing AI systems. Their guidelines aren't just theoretical - they're designed to be implemented in actual organizations dealing with real challenges.

> Additional information regarding the National Institute for Standards and Technology **found here**.

6.7. European Conformity - CE Regulatory Framework and Classification

The European Union Medical Device Regulation (MDR) establishes specific pathways for software as medical devices, including AI systems, through an integrated framework of MDR compliance, GDPR adherence, and harmonized standards (21). The regulatory process begins with classification under Rule 11 of Annex VIII, followed by conformity assessment procedures scaled to risk level.

GDPR integration requires documented data minimization strategies reducing personal data collection by at least 20% compared to baseline systems. Information Secuirty measures must achieve ISO 27001 certification, with penetration testing every six months.

Maintenance of CE marking requires MDR post market surveillance (PMS), including:

- Proportionate to the risk class of the solution, establish the PMS system
- Detailed post market surveillance plan, including:
    - Data associated with serious incidents and non-serious incidents, side effects, trend reporting, user feedback, and publicly available information about similar devices.
    - Data analysis for preventive or corrective actions
- Periodic safety update reports
    - Class IIa, IIb, and III devices require any changes to the benefit/risk profile, information on safety and performance, as well as results from post market clinical follow up
- Annual updates are required for high risk devices and semi annual for Class IIa devices

- Post Market Surveillance Report (PMSR) to be updated regularly and available upon request
- Systemic PMS monitoring for performance and emerging risk issues
- Data collected via PMS systems must be used to inform revisions to the monitoring system, advise regarding any manufacturing changes, and/or identifying corrective actions needed to improve safety or usability
- Relevant PMS data must be uploaded to the European Database on Medical Devices (EUDAMED) for regulatory oversight and transparency

Success in European markets demands rigorous adherence to these specific requirements while maintaining flexibility for evolving standards and clinical practice. This systematic approach ensures regulatory compliance while delivering sustained clinical benefit across intended use environments.

> More information regarding the European Conformity **found here**.
> This is an **European Conformity summary** provided by Perplexity.
> This is a **MDR summary** provided by Perplexity.

6.8. EU AI Act

The EU AI Act takes a comprehensive view of artificial intelligence, defining it broadly to include any systems that work with some degree of independence and produce outputs - whether those are content, predictions, recommendations, or decisions - that influence their operating environment. What makes this law distinctive, and where it differs from the approach outlined in HTI-1, is its focus on real-world impact rather than technical specifications. It's less concerned with how AI systems work internally and more interested in their effects on people and society.

The Act is careful to specify what isn't considered AI under its framework. This includes systems that simply follow human-defined rules automatically, basic traditional software, straightforward data processing that doesn't involve making inferences, and systems that lack autonomy. By excluding these, the Act maintains a clearer focus on truly autonomous systems.

At its heart, the Act creates a four-tier system based on risk levels: from systems with minimal or no risk, through those with limited risk, to high-risk systems, and finally to unacceptable

risk applications that are outright prohibited. This structured approach reflects the Act's broader goals of promoting transparency and safety while still allowing for innovation.
A central component of this Act is related to the technical documentation which is a requirement associated with its 'compliant by design' approach. This has particular focus on the higher risk strata of solutions with focus on transparency, documentation completeness, and life cycle coverage. Other key elements of its documentation requirements include:

- General Description
- Data Governance
- Algorithm and Model Information
- Risk Management
- Human Oversight
- Testing and Validation
- Logging and Monitoring

Importantly, the Act isn't starting from scratch. There are already examples of AI systems successfully operating under other regulatory frameworks, particularly those with CE marking. These existing implementations can serve as valuable models for how to meet the Act's compliance requirements effectively.

More information regarding the EU AI ACT is **found here**.
Also, this is an **EU AI Act summary** provided by Perplexity.

6.9. TEFCA - The Trusted Exchange Framework and Common Agreement

TEFCA (Trusted Exchange Framework and Common Agreement) represents a significant shift in healthcare data sharing that emerged from HHS in 2023. While it doesn't specifically address artificial intelligence, this framework could fundamentally reshape how healthcare data moves across the country, which could either affect how one gains access to the critical data on which AI needs to be trained and/or the multisite/multimodal workflows it may be able to enable.

At its core, TEFCA builds upon existing health IT standards but introduces the Qualified Health Information Network (QHIN) architecture. This new approach to data exchange creates a standardized way for healthcare information to flow between networks, complete with centralized directories, consistent response patterns, and strict timing requirements for data sharing.

This new framework has important implications for AI in healthcare. For AI tools to remain practical in clinical settings, especially in time-sensitive situations, they'll need to be capable of handling the rapid cross-network data exchange that TEFCA enables. Even the most accurate and technically sophisticated AI tools could become impractical at scale if they can't integrate with TEFCA's network structure. An example to contemplate for this would be in the instance when one had a very capable but very large model which, in order to properly function, needs complete access to appropriate patient information which may be stored within disparate databases within disparate health systems (domains). As this model is directed to access and analyze the required information it will need to do so in a QHIN compliant manner while still satisfying the accuracy and timeliness requirements of the given use case.

Looking ahead, AI tools will need to adapt to QHIN data exchange patterns to maintain their utility in clinical settings. While existing AI solutions can continue functioning through local integration, cached data approaches, or institutional workarounds, these are likely temporary solutions. For AI to achieve widespread, industry-wide adoption in healthcare, it will need to evolve alongside TEFCA's information exchange architecture.

More information regarding the Trusted Exchange Framework and Common Agreement **found here**.
Also, this is an **TEFCA summary** provided by Perplexity.

7. Epilogue: A Standard

Gemini Advanced 2.0 Flash rendering of a Standard Oil kerosene can.

It was from extremely humble beginnings in Cleveland, Ohio when over 150 years ago John D. Rockefeller started the Standard Oil Company in an effort, not to literally create the oil and gas industry as we all still know it today, but rather to provide a stable, standard means for people to light their homes. Standard Oil was a lighting company as kerosene was the most common method to bring light to darkness after the sunset each day. From this nexus, business itself was transformed as the means of transportation, branding, merger/acquisition, oil exploration, industrialization, and the legal framework under which all commerce operates today was set on a completely new trajectory. All so we could, for instance, see one another's faces at an evening dinner table.

Our goal is to establish a standard, and we emphasize that developing an AI model is only the beginning of what's required for production. BRIDGE has endeavored to make very clear the required capacity of any resource (academic institution, private vendor, or individual contributor) to successfully achieve adoption is to deliver a complete solution. That solution must include the first, the last, and all the steps in between in order for one's targeted user to fit this novel instrument into their daily grind - and healthcare today is a monstrous grind. It needs a relief valve. It needs a standard in order to induce an entire vertical of innovation to begin to leverage its collective energies in a reproducible manner.

It is our contention that clinical AI solutions are and will continue to play a vital role in delivering the relief craved by so many clinicians today. By working together and with a standard set of expectations as to how this deliverance can be achieved, we may begin to experience a collective acceleration of trust in clinical AI technology as those expectations are met.

It is with humility with which we acknowledge that, just as it was for JD, it is the consumer who determines if something can become a standard. It is your choice. Our part is to publish the above as one possible means to begin to achieve it. While we simultaneously acknowledge no single entity is fully in control of this pursuit, we are also not alone in the endeavor. If we accomplish this esteemed goal, we will do it together.

Appendix

A.1. Work Cited

1. Callahan, Alison. “Standing on FURM ground - A framework for evaluating Fair, Useful, and Reliable AI Models in healthcare systems” **NEJM**, September 18, 2024, vol 5; No 10
2. Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices **https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices**
3. Model Legislative Language – AI Qualified End-User, American College of Radiology, October 2024.
4. Langlotz, Curtis. “Will Artificial Intelligence Replace Radiologists?” **https://pubs.rsna.org/doi/10.1148/ryai.2019190058.**
5. Wong, Andrew, et al, “External Validation of a Widely Implemented Proprietary Sepsis Prediction Model in Hospitalized Patients” **JAMA**, June 21, 2021
6. Maier-Hein, L., Reinke, A., Godau, P. et al. Metrics reloaded: recommendations for image analysis validation. Nat Methods 21, 195–212 (2024) **https://doi.org/10.1038/s41592-023-02151-z**
7. Solutions Architect Salary in USA and India 2025 [Earnings Revealed], Renee Jones, January 30, 2025, **https://www.netcomlearning.com/blog/solutions-architect-salary**
8. Wiklund, Peder, et al, 2022. Incidental pulmonary embolism in patients with cancer: prevalence, underdiagnosis and evaluation of an AI algorithm for automatic detection of pulmonary embolism. ***Springer Nature***, August 25, 2022, Volume 33, pages 1185–1193
9. Armoundas, Antonis, et al, 2024. Use of Artificial Intelligence in Improving Outcomes in Heart Disease: A Scientific Statement From the American Heart Association. ***Circulation***, Volume 139, Number 14
10. Petry, Michael, et al, 2022. Decreased Hospital Length of Stay for ICH and PE after Adoption of an Artificial Intelligence-Augmented Radiological Worklist Triage System. ***Radiology Research and Practice***, August 2022
11. Guardado, J.R., 2017. Policy research perspectives. *Chicago, IL: American Medical Association.*
12. Seabury, S.A., Chandra, A., Lakdawalla, D.N. and Jena, A.B., 2013. On average, physicians spend nearly 11 percent of their 40-year careers with an open, unresolved malpractice claim. ***Health Affairs***, January 2013, 32(1), pp.111-119.
13. Banja, J.D., Hollstein, R.D. and Bruno, M.A., 2022. When artificial intelligence models surpass physician performance: medical malpractice liability in an era of advanced artificial intelligence. ***Journal of the American College of Radiology***, July 2022, 19(7), pp.816-820.
14. Song, E.C., Bernstein, M.H., Lay, P.S., Druart, L., Dibble, E.H., Lourenco, A.P. and Baird, G.L., 2024. Accessing AI mammography reports impacts patient interest in pursuing a medical malpractice claim: The unintended consequences of including AI in patient portals. ***medRxiv***, December 29, 2024, pp.2024-12.
15. Bernstein, M.H., Sheppard, B., Bruno, M.A., Lay, P.S. and Baird, G.L., 2024. Just because you’re paranoid doesn’t mean they won’t side with the plaintiff: Examining perceptions of liability about AI in radiology. ***medRxiv***, August 1st, 2024, pp.2024-07.
16. Stempniak, Marty, 2023. Radiology Partners reaches ‘milestone,’ deploying clinical AI across more than 20 million exams. ***Radiology Business***, November 13, 2023.
17. Newman-Toker DE, Nassery N, Schaffer AC, et al Burden of serious harms from diagnostic error in the USA, ***BMJ Quality & Safety***, January 19, 2024.
18. Software as a Medical Device **https://www.fda.gov/medical-devices/digital-health-center-excellence/software-medical-device-samd**
19. Secure Software Development Practices for Generative AI and Dual-Use Foundation Models, NIST Special Publication 800, NIST SP 800-218A, July 2024 **https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-218A.pdf**
20. AI Risk Management Framework, January 2023 **https://www.nist.gov/itl/ai-risk-management-framework**
21. Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices **https://eur-lex.europa.eu/eli/reg/2017/745/oj/eng**