How to Generate Clinical Evidence For Artificial Intelligence.

An expert guide by:

Prova Health

“

AI is a Tool.

The choice about how it gets deployed is Ours.

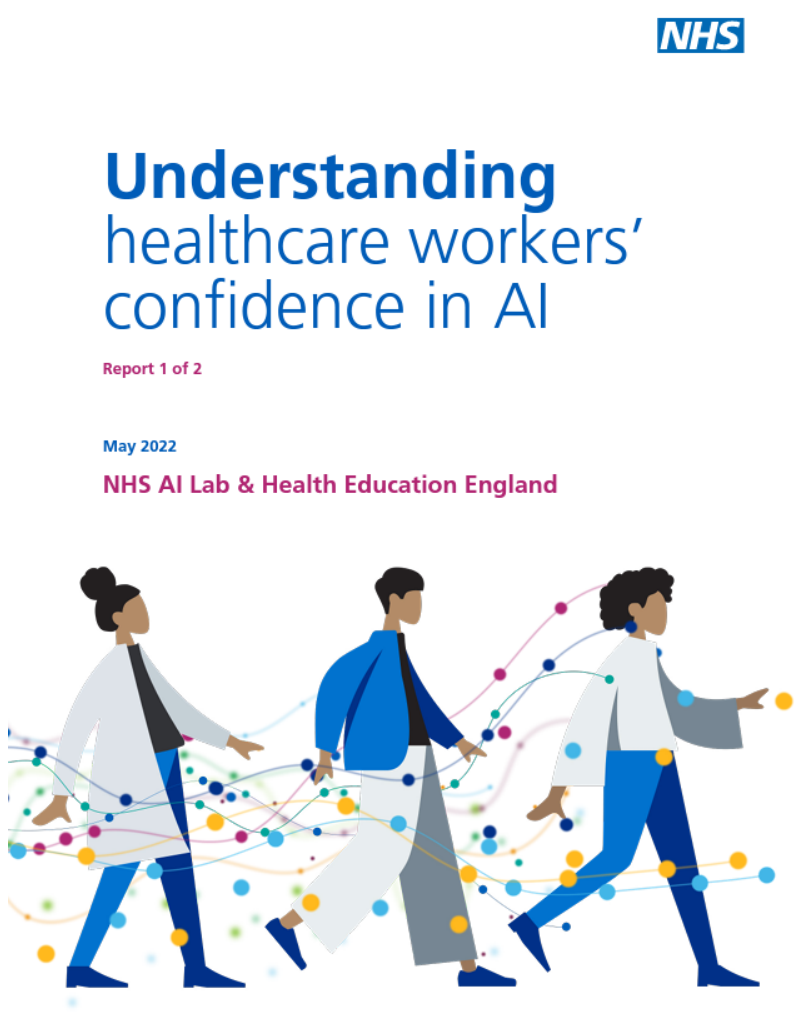
Novel AI solutions are being created to address some of the biggest challenges in the **prognosis, diagnosis, and treatment of disease**, as well as clinician workflows and service improvement.

However, adoption of digital solutions and AI in healthcare is slower than in other industries.

The majority of clinicians don't have direct experience with AI technologies.

Only a quarter have recommended a digital therapeutic, and **less than** a **fifth** have prescribed one.

Clinicians' perspectives on AI



A recent study by Health Education England found that **clinician confidence in AI is largely dependent** on how AI solutions are governed.

It highlighted the need for a robust, **AI-specific medical device regulation pathway**, and guidance on the safe and effective use of AI tools.

Clinicians expect **sufficient levels of regulation and robust validation** of AI-based medical devices.

Clinician Concerns

- Lack of test cases for regulatory approval.
- Lack of clinical guidelines.
- Lack of generalisability to target population
- Cost to healthcare system and demonstrating cost effectiveness
- Challenges in system integration and sustainability
- Data privacy.

While a majority (68%) of clinicians are excited about the potential of AI in healthcare, less than a third have used it in practice.

Their key concerns are a **lack of training** (62%), **doubts about efficacy** (48%), and **a lack of clinical evidence** validating these tools (45%).

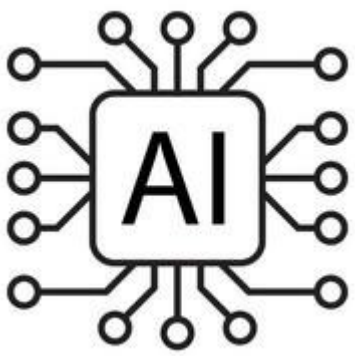
Evidence for AI solutions.

“

As of 2022, 44% of the leading digital health companies in the USA had no regulatory filings or published clinical trials for their solution

(Day et al., 2022).

Unlike other types of digital health solutions, for AI products **there are two levels of evidence** and validation:

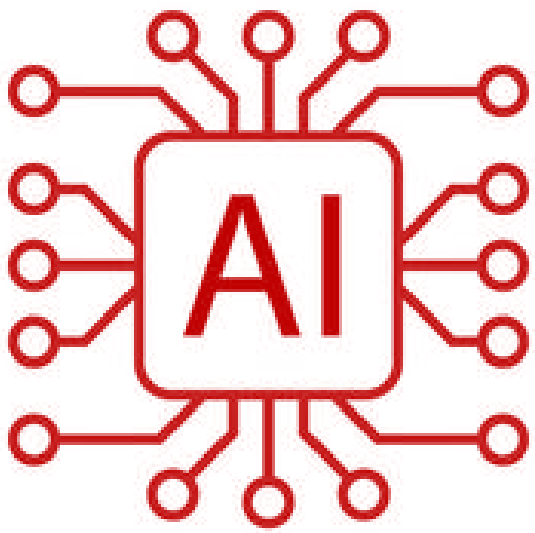


The Algorithm
(“model evidence”)

The Product in which the algorithm is embedded.
(“solution evidence”).

Once a model has been internally and externally validated from a data perspective, **the solution as a whole needs to be evaluated**.

Clinical pathways need to be **scrutinised pre and post implementation of a solution** to ensure that any improvements are actually being captured.



Model evidence consists of internal and external validation.

An innovator will develop a model from a given data set and conduct internal validation.

This will indicate how accurate and reliable the model is, but will be **limited to the dataset that the client has access to.**

Testing the model on an external dataset (external validation) tests for how well the model performs on other datasets otherwise known as '**generalisability**'.

AI EVIDENCE ROADMAP

EARLY VALIDATION

Internal Model Validation

Model tested with **small fictional set** of patient cases.

ACCURACY MODEL BIAS

External Model Validation

Multiple rounds of testing with a **large set of synthetic patient cases** based on real-world EHR data.

USER BIAS BEHAVIOURAL CHANGE EARLY HEALTH ECONOMIC OUTCOMES

Clinical Simulation

Remote simulation study with **100 clinician triaging 50 high fidelity patient cases** each.

SAFETY WORKFORCE IMPACT OPERATIONAL EFFICIENCY

Real World Trial

Real world observational study in a group of GP practices evaluating end to end patient journey for accessing GP appointments.

IN SIMPLE TERMS,

Developers of AI solutions (and those wishing to adopt them) should **answer with robust evidence** at least the following questions:

- Is the solution addressing **a real clinical or operational problem**?
- Does the model perform well on the **developer's own datasets?**
- Does the model perform well **to external datasets?**
- Does the solution that includes the model **address the problem effectively?**
- More specifically, does the solution **reduce or eliminate existing biases and inequalities**?
- Does the solution **deliver value for money** to the client?

Guideline	Framework	Area
CONSORT-AI	Consolidated Standards of Reporting Trials-AI	Reporting guideline
DECIDE-AI	Developmental and Exploratory Clinical Investigation of DEcision-support systems driven by Artificial Intelligence	Clinical Decision Support
PROBAST-AI	Prediction model Risk Of Bias ASsessment Tool-AI	Diagnostic and prediction models
QUADAS-AI	QUality Assessment tool for artificial intelligence-centred Diagnostic test Accuracy Studies	Diagnostic accuracy
SPIRIT-AI	Standard Protocol Items: Recommendations for Interventional Trials-AI	Reporting guideline
STANDING Together	STANdards for Data INclusivity and Generalisability	Representative data
STARD-AI	Standards for Reporting of Diagnostic Accuracy Study-AI	Diagnostic accuracy
TRIPOD-AI	Transparent Reporting of a multivariable prediction model of Individual Prognosis Or Diagnosis-Artificial Intelligence	Prediction models

Table 1: AI evaluation and clinical trial reporting guidelines and tools[12]

Various frameworks have been developed to help standardise expectations and reporting outcomes for specific types of AI technologies.

These are helpful for developers building a product **that matches one of these use cases**.

CURRENT CHALLENGES.

- **AI models can be considered a "black box"** due to difficulty in understanding how the model is producing an output.

- **Lack of sufficient data to train AI models on**, or ensuring the data is from a population that is representative of that in which the solution is going to be implemented

- AI systems rely heavily on large datasets, and ensuring that **patient information remains confidential and secure** is paramount.

- Breaches in data privacy and security can erode patient trust and **have significant legal implications**.

- **Logistical difficulties in implementation**, and consideration of the barriers to adoption as well as of the necessary sociocultural or pathway changes.

Evidence is required across the product lifecycle.

Certain types of evidence are more strongly associated with **specific stages of the product life cycle.** For example:

- **Secondary research**

(reviewing existing research, which helps innovators to better understand a clinical problem)

- **User research**

(which can help validate a solution concept)

- **A/B testing**

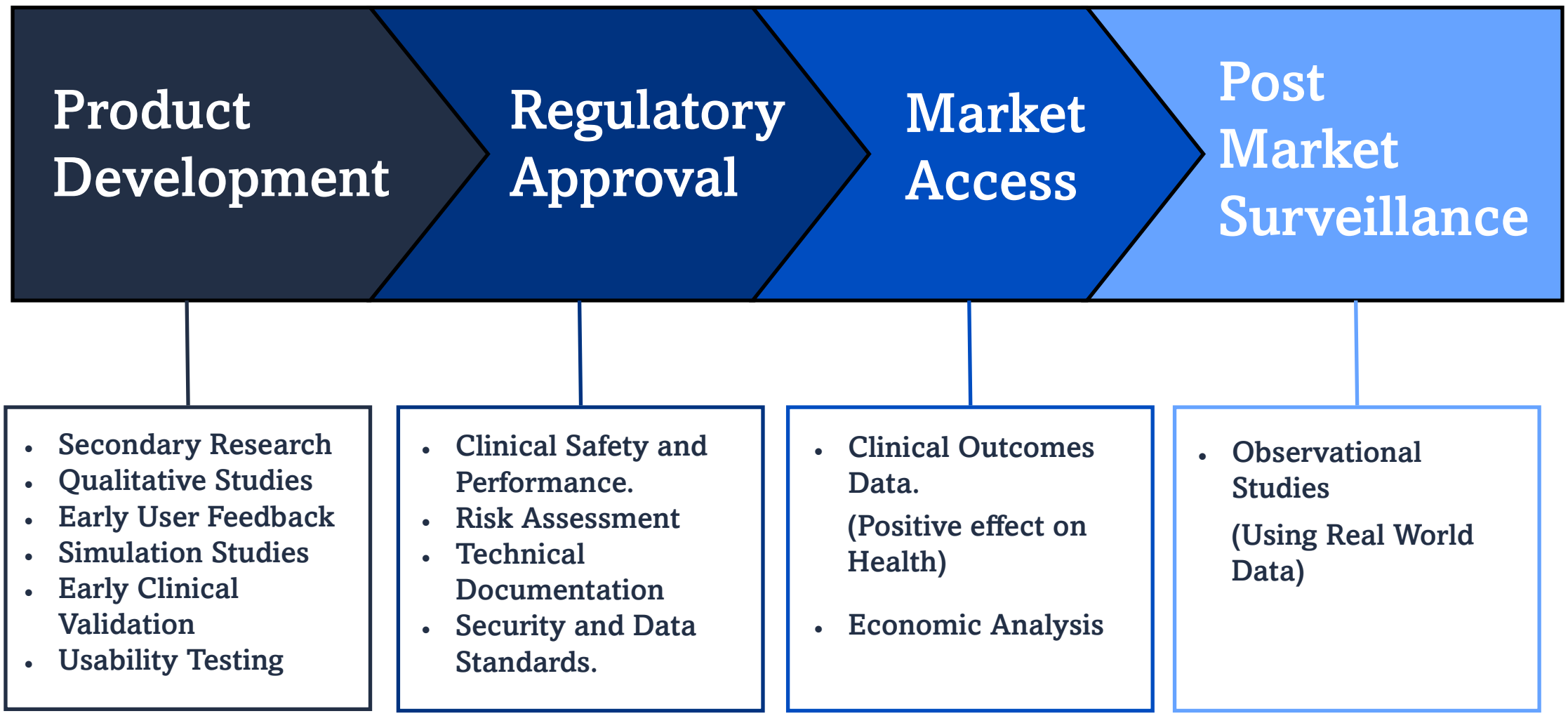
(which allows comparison of different versions or features).

- **Clinical data**

Demonstrating safety and clinical performance may be critical for regulatory certification.

- **Economic analyses**

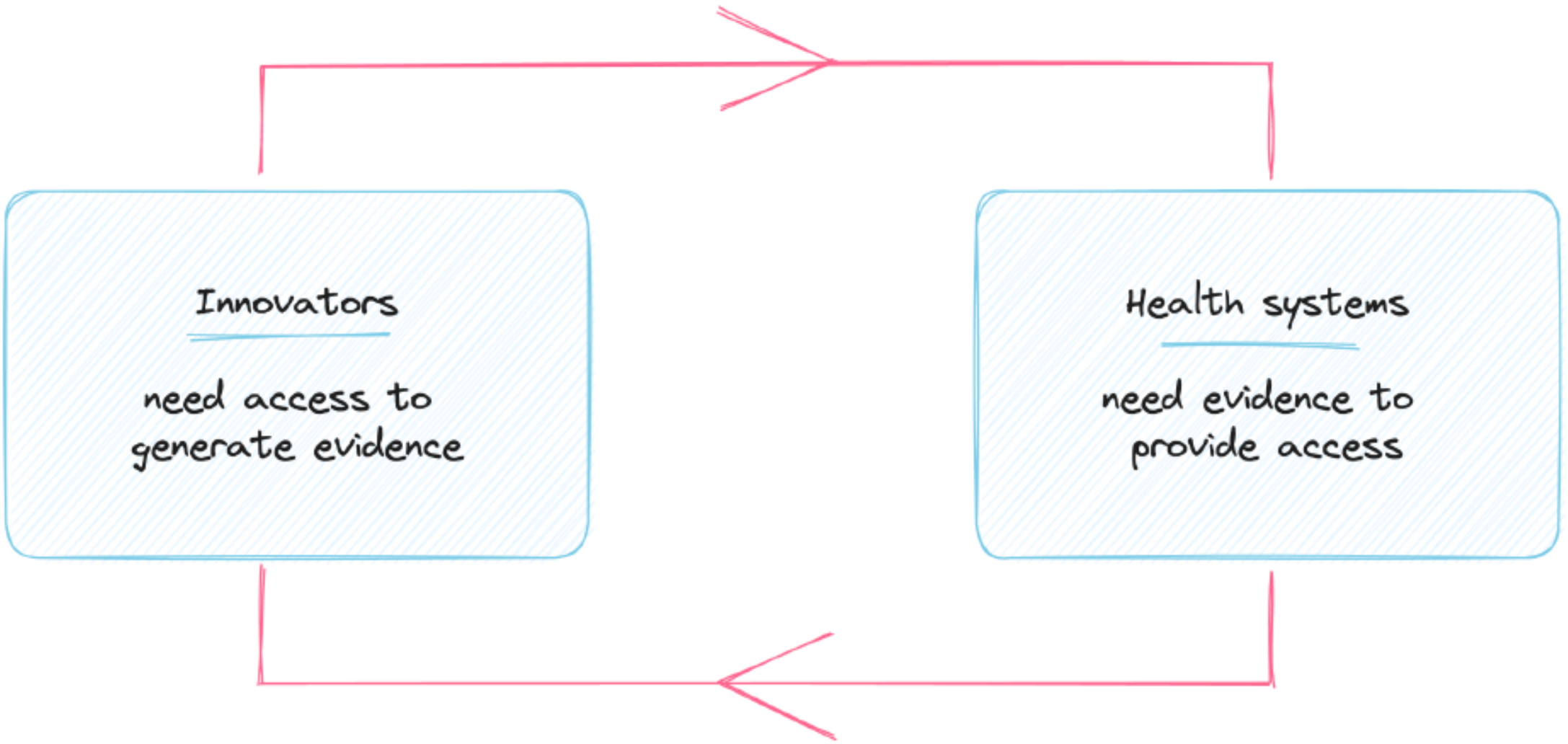
Important for showing evidence of value to health systems in order to sell a solution.



Examples of Evidence Generation at different stages of the product life cycle.

Overcoming the Evidence Limbo.

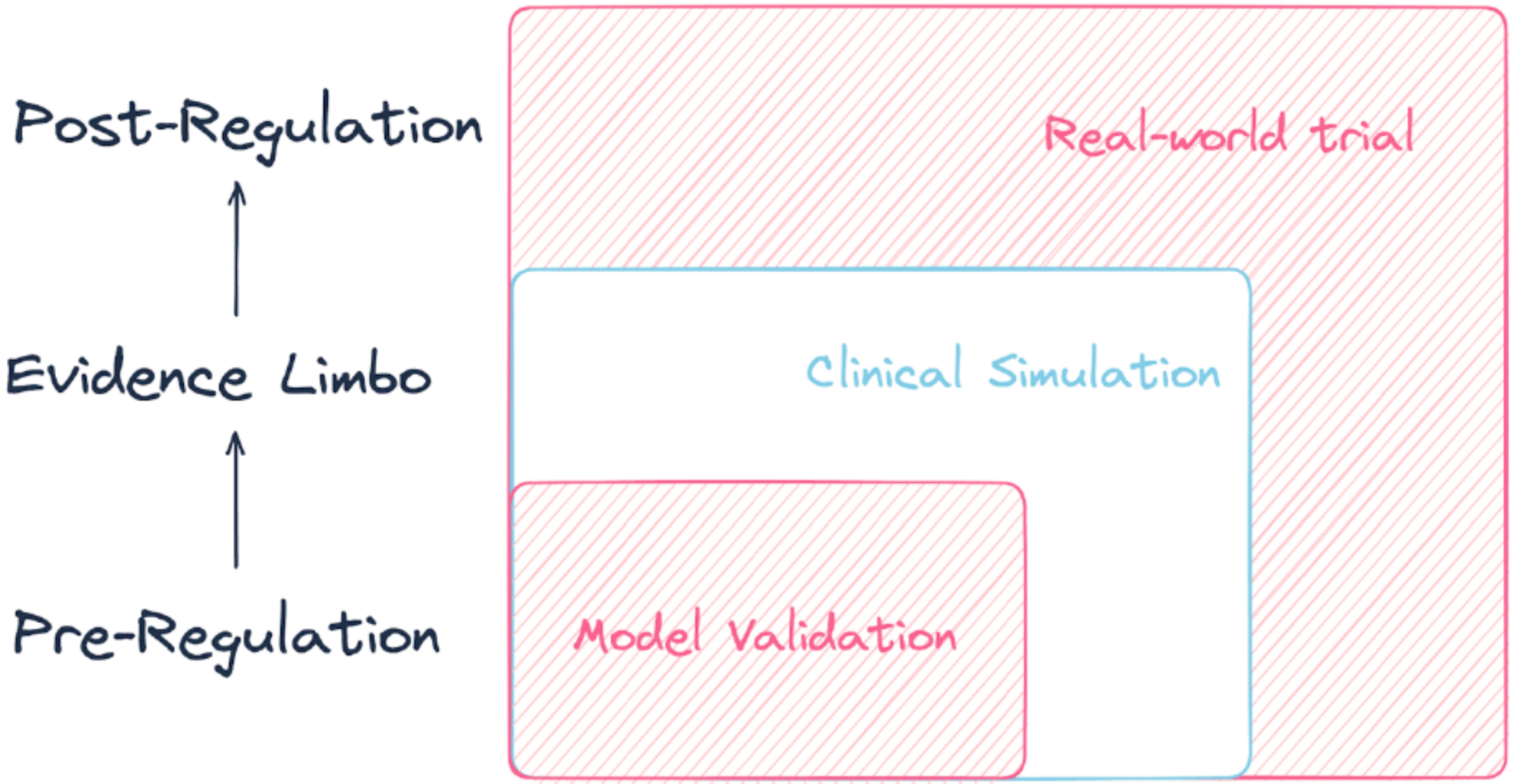
THE EVIDENCE LIMBO



External model validation can be difficult without access to good external datasets.

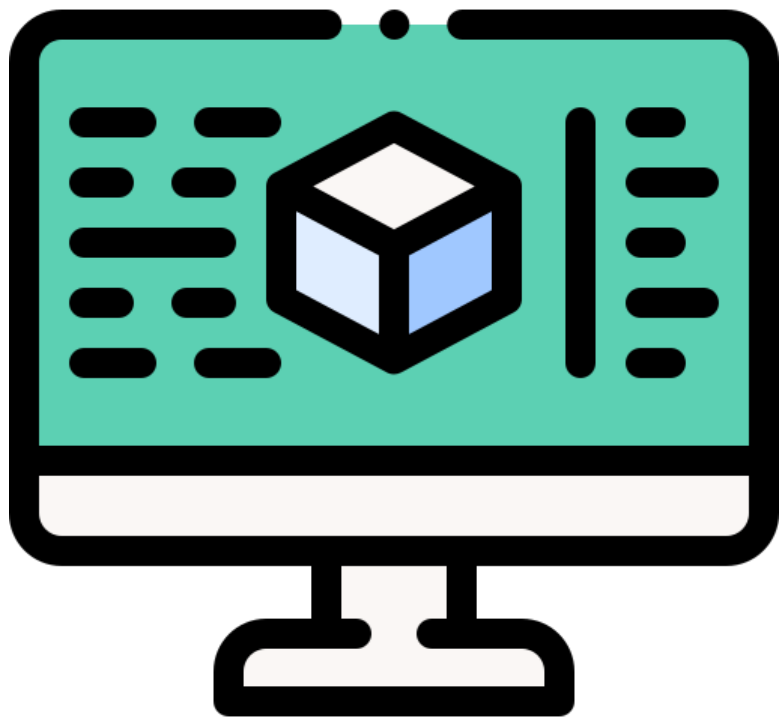
Such challenges become even more significant when innovators seek to enter the market and generate real-world solution evidence.

To achieve deployment, they may be **required to generate evidence before their solution is actually adopted in the real world**.



Thankfully, innovators can rely on innovative methodologies like **clinical simulation** to generate evidence in a way that addresses the “evidence limbo”.

Simulation has emerged as a promising tool that enables safe, efficient and cost-effective evaluation of digital health solutions in a controlled environment.



Clinical Simulation is achieved by closely replicating their intended real-world use in a controlled environment.

Remote, multi-site trials can be conducted at relatively low cost using virtual communication platforms.

Simulation studies are highly scalable and flexible and study designs can easily be adapted to keep up with the frequent updates to digital solutions.

BENEFITS OF SIMULATION

Ability to use synthetic patient data.

Realistic, synthetic datasets can be modelled using real data in a way that minimises privacy concerns while preserving the complexities of the data.

Include higher risk patients

Simulation studies also allow researchers to test solutions with **data representative of higher risk patients**, who are often excluded.

Reduce Biases

Allows for more extensive testing of subpopulation data helping to **alleviate the risk of biases** (e.g., ethnicity, gender)

Study human-AI interaction

Help better understand and study the human-AI interaction and the potential for AI to impact on **decision making and clinical behaviours**.

Ultimately,

Evidence generated in simulation studies is **unlikely to be sufficient on its own** to support decisions around regulatory approval for higher-risk solutions.

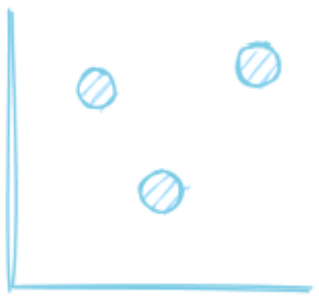
However, it is a pragmatic adjunct to established methods that can be used to generate evidence of **reasonable strength.**

Real World Evidence.

The majority of AI studies have been **retrospective** in more tightly controlled conditions and have relied on comparing clinical expert performance vs algorithm performance.

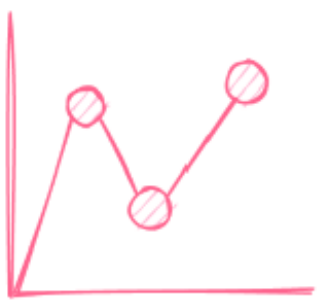
Ideally an AI tools performance should be compared to the performance of a pathway **pre-implementation**.

Real world studies relying on **prospective data are key** to informing the true clinical utility of an AI solution in practice in any given workflow.



Real-world data (RWD)

= data relating to patient health status and/or the delivery of health care routinely collected data form a variety of sources



Real-world evidence (RWE)

= clinical evidence about the usage and potential benefits or risks of a medical product derived from analysis of RWD

In comparison with randomised controlled trials (RCTs), which may inform researchers about how an intervention performs for a specific group (under tightly controlled conditions),

RWE can provide more certainty about how effective a solution is when deployed in the real world.

These studies can also illuminate potential unintended consequences of technologies on different population groups, for example minority ethnic groups, identifying potential bias and other negative outcomes, which can then be addressed in a timely manner.

Medicines &
Healthcare products
Regulatory Agency

To note,

Regulators in several countries have begun to **encourage the use of RWE to inform regulatory decisions** in the post market phase.

Economic Evidence

Another key piece of evidence required for AI solutions is **economic evidence.**

As AI solutions are being launched to the market at a rapid speed, there is **frequently insufficient data** to support their efficacy.

Traditional health technology assessment (HTA) approaches, which rely on published research, **can be time-consuming and may not be compatible** with the quick development cycles of digital health technologies (DHTs).

Economic evidence is critical because it assesses **the cost-effectiveness** and **economic impact** of these technologies especially in health systems with constrained resources.

This evidence will assist decision makers in determining **if investing in a dedicated solution will provide value** in terms of better patient outcomes relative to the expenditure involved and **in comparison to the incumbent pathway.**

Hope you found this helpful!

This is a series we are making to help HealthTech Innovators access better resources.

Just our small way of helping!