Heart Failure Prediction

GROUP #6

Preamble

How it all began from the beginning

Preamble

Source:

Before we did our XAI, we had heard several daily comments from some people who tend to ask for help and even inquire an advice regarding on health issues.

Discovery:

When we researched this, we observed that people (hereafter referred to as patients) tend to require consultations with doctors for check-ups, they are often routine ones.

What problem was there in this context that made it a routine need?

Interview

We requested an interview with some patients, doctors and even data scientists

Acknowledgments to the interviewees
from their feedback and time

Interview

Our volunteers:

We selected 3 patients to tell us their troubles and habits. From this points we requested interviews (quite hard step) from 3 doctors and 3 data scientists after their working time.

What did we found out:

There were very varied and complex reasons, some of which we cannot implement or contemplate, but many of them we believe we can cover the most common and criticized situations in the interviews.

What user histories are translated from their goals, habits and frustations?

Patients

Doctors

DataSc.

Don't care about medical terms, they prefer something that is simple and within his/her reach [6]

They support the idea that explanatory AI should present potential risks even if they are uncertain [6]

Concerns about sending/breach - - data over the Internet [5]

They want to see a breakdown of risks and their intensities at first glance [4]

Prefers to visually display traffic light colors in the verdict [3]

Make inquiries on your own and decrease routine doctor visits [2]

Concerned about biased data model/prediction [3]

Not clear about how to interpret blood pressure [2]

Using SHAP visualization [2]

Using LIME visualization [1]

10 user histories

Related works

There were already existing?
something similar…?

Related works

HeartScore: https://heartscore.escardio.org/Calculate/quickcalculator.aspx?model=low

Calculates the cardiovascular risks based on Systolic and Cholesterol with tendencies after 5 years; however half of these fields are medical/doctor exclusive, not for patients.

Calculadora de Riesgo Cardiovascular: https://fundaciondelcorazon.com/prevencion/calculadoras-nutricion/riesgo-cardiovascular.html

A plain calculator, is very simple and likely to be used by patients, however it doesn't respect the user choices, such as telling “not smoking” when the patient isn’t smoker.

Which our user stories is already covered by both works? (minigame)

Patients

Doctors

DataSc.

Don't care about medical terms, they prefer something that is simple and within his/her reach [6]

They support the idea that explanatory AI should present potential risks even if they are uncertain [6]

Concerns about sending/breach - - data over the Internet [5]

They want to see a breakdown of risks and their intensities at first glance [4]

Prefers to visually display traffic light colors in the verdict [3]

Make inquiries on your own and decrease routine doctor visits [2]

Concerned about biased data model/prediction [3]

Using SHAP visualization [2]

Not clear about how to interpret blood pressure [2]

Using LIME visualization [1]

? of 10

CAN YOU GUESS WHICH ONES ARE?

Patients

Doctors

DataSc.

Don't care about medical terms, they prefer something that is simple and within his/her reach [6]

They support the idea that explanatory AI should present potential risks even if they are uncertain [6]

Concerns about sending/breach - - data over the Internet [5]

They want to see a breakdown of risks and their intensities at first glance [4]

Prefers to visually display traffic light colors in the verdict [3]

Make inquiries on your own and decrease routine doctor visits [2]

Not clear about how to interpret blood pressure [2]

Concerned about biased data model/prediction [3]

Using SHAP visualization [2]

Using LIME visualization [1]

1 of 10 user histories

Patients

Doctors

DataSc.

Prefers to visually display traffic light colors in the verdict [3]

1 of 10
user histories

:(

You heart ran into a problem and needs to restart. We're just collecting some error info, and then we'll restart for you.

If you call a support person, give them this info:

Prefers to visually display traffic light colors in the verdict [3]

Related works

…Just “*one*”?
Definitely there’s a gap,
here comes our XAI

Interface design

Patients:

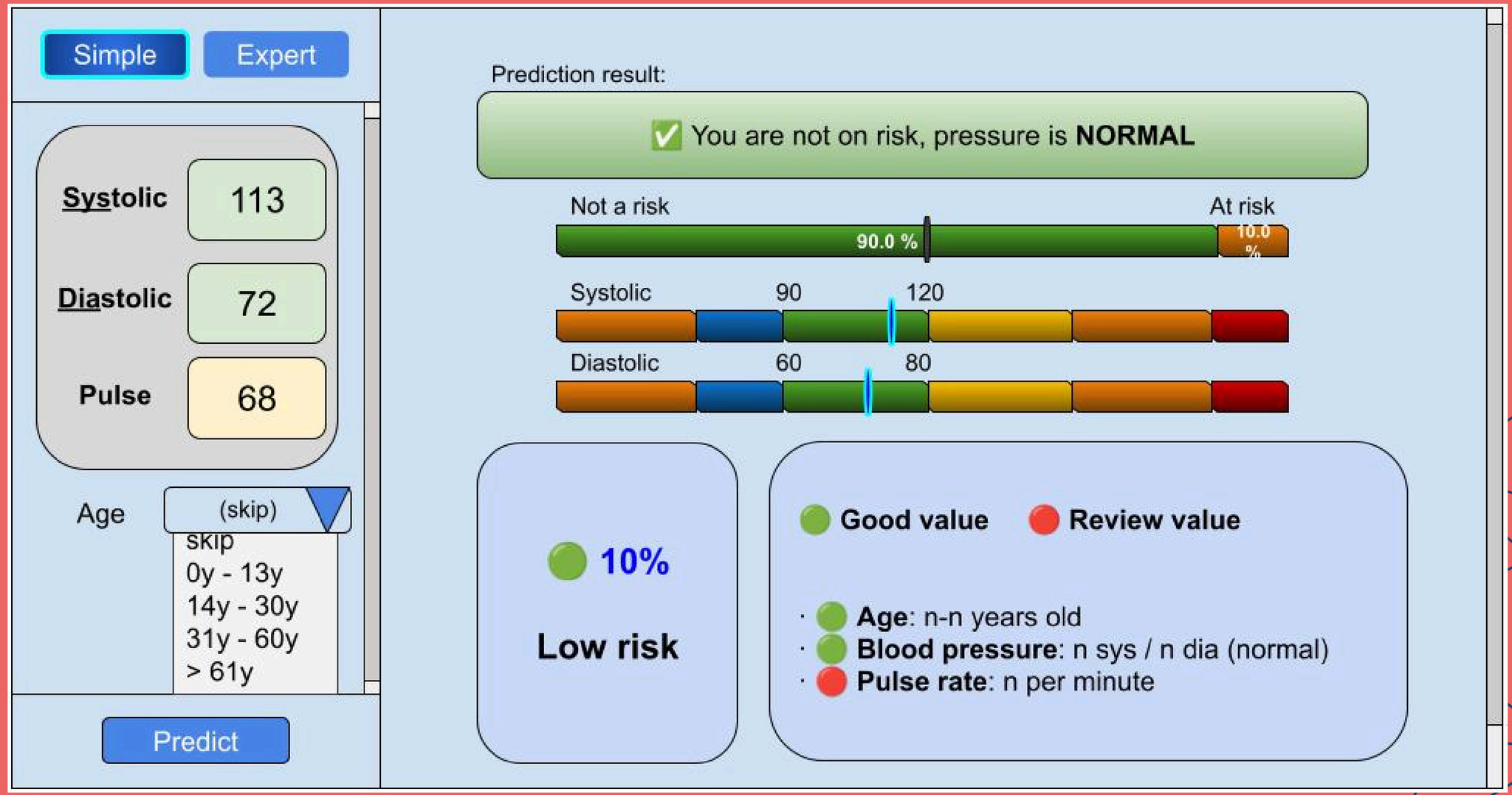
It must have a very **simple** fields, does **not use medical exclusive** equipment/monitor, **can self-check** itself without requiring doctor's/friends supervision, **offline usage**

Doctors:

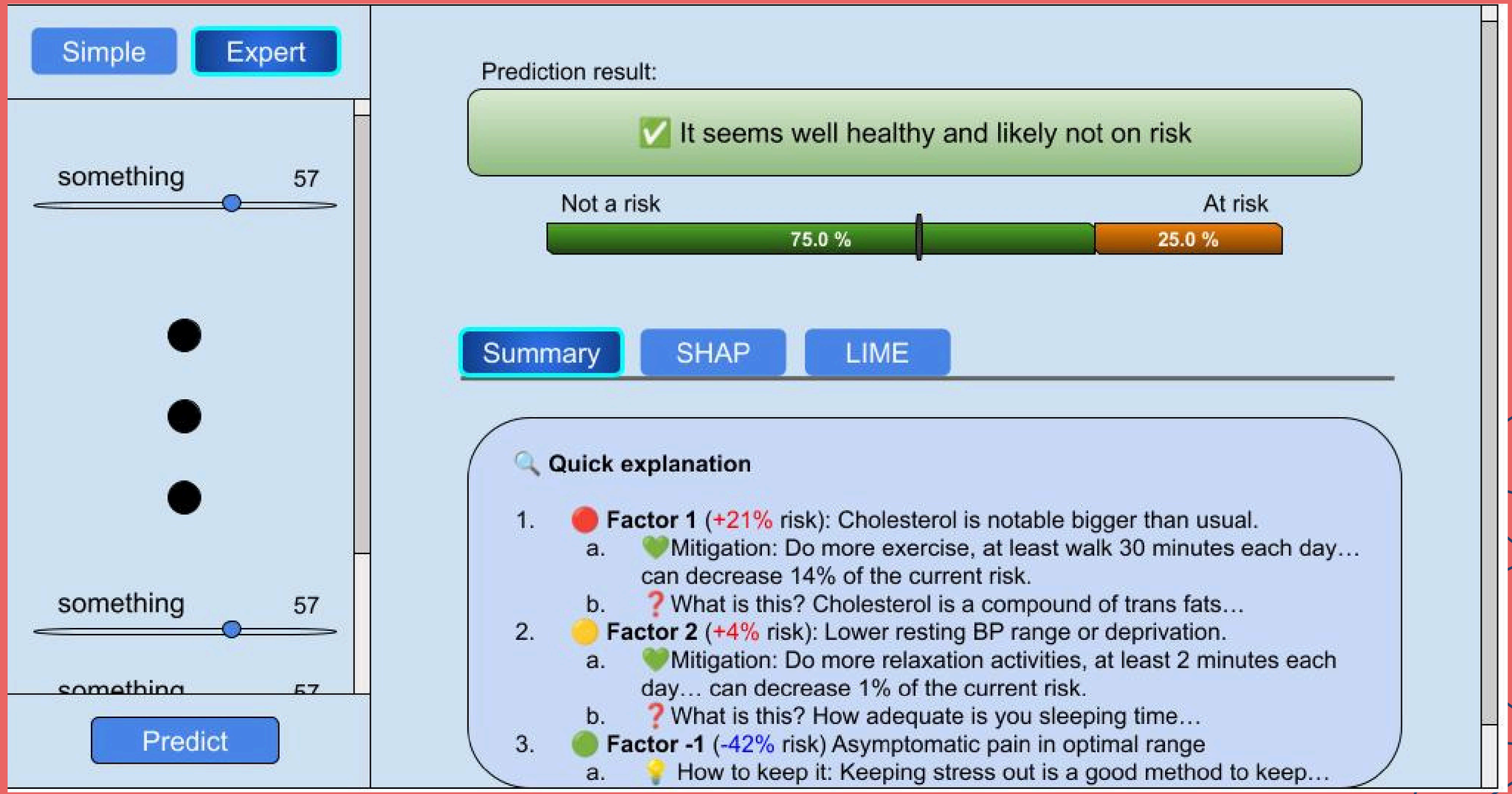
It must have a very **detailed** fields, **with medical exclusive** equipment/monitor, **onsite check-it** with supervision, representation **with SHAP + LIME** describing the results

How our users will interact with our XAI?

Interface design (patients)



Interface design (doctors)



Interface design

Now we have a clear idea from our users, the interface and interactions

Later there will be a live demo!

Project development

What data will we use?

Project development

Key milestones in our heart failure prediction system journey

Dataset

Analyzed heart failure data: 918 patients, 11 features

Model Selection

Chose Random Forest for interpretability and clinical reliability

Training

Achieved 89% accuracy with 92% disease detection recall

XAI Implementation

Integrated SHAP and LIME for transparent decision explanations

Streamlit Web

Built dual-interface dashboard for patients and doctors

Dataset Characteristics

Characteristics:

- 918 patients from Kaggle Heart Failure Prediction dataset
- Balanced distribution with 55% positive cases and zero missing values
- Features span age (28-77), blood pressure, cholesterol, ECG results, and exercise-induced symptoms

Insights:

- Strong correlations identified: Oldpeak, FastingBS, and Age with heart disease risk
- Feature preprocessing included standardization and one-hot encoding

Model Selection Decision:

- Random Forest chosen for robust performance and XAI compatibility
- Perfect foundation for SHAP and LIME explanations required in healthcare decision-making contexts

Model Training & Performance

Training & Validation Process:

- **Random Forest model** trained on 80% of patient data (719 cases) with balanced treatment of healthy and at-risk patients
- **Data preparation** included scaling numerical values and converting categories to numbers
- **Rigorous testing** on separate 20% of data (180 patients) to ensure reliable performance on unseen cases

Classification Report:

	precision	recall	f1-score	support
0	0.90	0.85	0.87	81
1	0.88	0.92	0.90	99
accuracy			0.89	180
macro avg	0.89	0.89	0.89	180
weighted avg	0.89	0.89	0.89	180

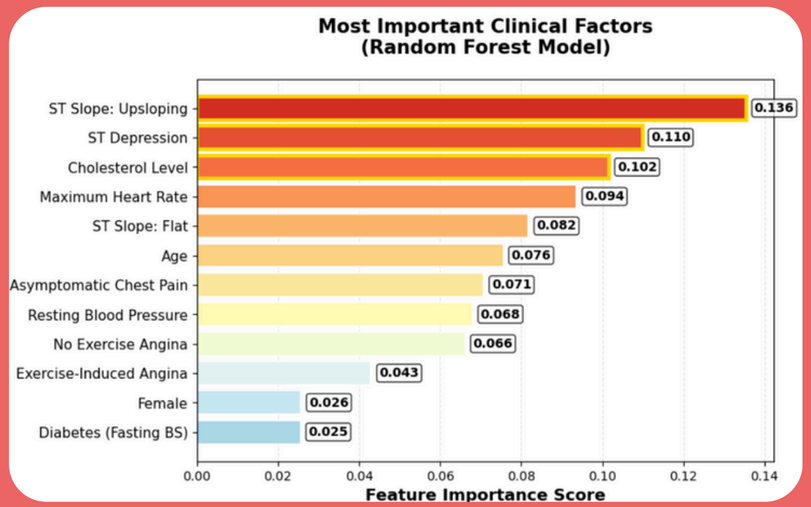
Performance Results:

- **89% accuracy overall** with strong performance on both groups: correctly identifies 90% of healthy patients and 88% of at-risk patients
- **Excellent disease detection**: captures 92% of truly at-risk patients, meaning only 8 out of 100 high-risk cases would be missed
- **Strong clinical reliability:** out of 180 test patients, correctly identified 81 healthy cases and 99 disease cases

SHAP Explanations

What is SHAP and How It Helps:

- **SHAP (SHapley Additive exPlanations)** quantifies each feature's contribution to every prediction, showing which medical factors push risk up or down
- **Global model understanding** reveals patterns across all patients: how age, blood pressure, cholesterol, and other factors generally influence heart disease risk



Our SHAP Results:

- **Key risk factors identified**: ST_Slope features, FastingBS (diabetes), and Cholesterol emerge as most influential across patient population
- **Feature interaction patterns** show how combinations like "flat ST slope + high cholesterol" significantly increase risk predictions

LIME Explanations

What is LIME and How It Helps:

- **LIME (Local Interpretable Model-agnostic Explanations)** creates simple explanations for individual patient predictions by learning locally around each specific case
- **Personalized insights answer** "Why did THIS patient receive THIS risk score?" by showing exact feature contributions for individual cases
- **Local approximation** builds interpretable linear models around each patient to reveal which specific factors drove their personal risk assessment

Our SHAP Results:

- **Three patient examples:** Low-risk patient protected by flat ST slope and asymptomatic pain, while moderate and balanced-risk patients show upsloping ST patterns as primary risk drivers with varying protective factors.

Streamlit Implementation

Why Streamlit:

- **No web development needed** - unlike Flask or Django that require HTML/CSS knowledge, Streamlit lets us build web apps using only Python code we already know
- **Medical-focused simplicity -** while Gradio or Dash were options, Streamlit's clean interface and professional appearance fits perfectly with healthcare environments
- **Built-in interactive widgets** - automatic sliders, dropdowns, and buttons that work instantly without complex coding, ideal for clinical parameter input

Model Integration:

- **Smart model loading -** the app loads our trained model file once when started, then reuses it for all patients instead of reloading every time (faster performance)
- **Identical data processing -** user inputs go through the exact same cleaning and formatting steps we used during training, ensuring reliable predictions
- **Consistent explanations** - SHAP and LIME use the same data preparation, so explanations match what the model actually learned

Simple vs Expert Mode

User-Centered Design :

- **Clear user needs identified -** interviews with 9 people showed patients want simple language and basic questions while doctors need precise medical measurements
- **Patient-friendly approach -** Simple mode uses everyday terms like "age range 50-59" and "diabetes yes/no" instead of confusing medical jargon that scared patients away

Implementation Strategy:

- **One toggle, two experiences -** flipping a single switch changes everything: Simple mode asks 5 basic questions, Expert mode shows 11 detailed medical sliders
- **Different calculation methods -** Simple mode adds up points (older age = more points, high blood pressure = more points) while Expert mode uses the full AI model directly

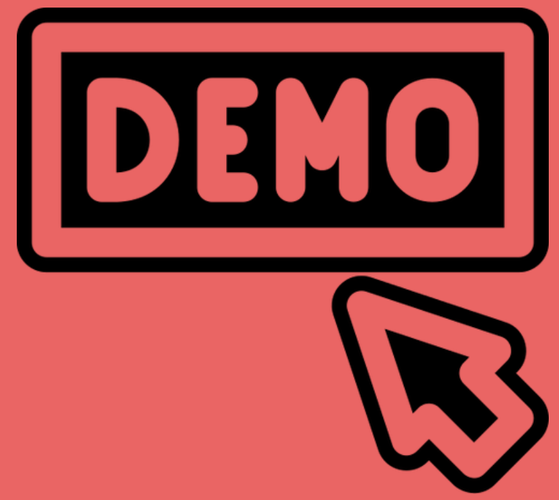
Results:

- **Patients succeeded independently -** after struggling in Round 1, patients easily used Simple mode alone in Round 2 testing with no help needed
- **Doctors maintained precision -** Expert mode gave healthcare professionals the detailed control they demanded while still providing clear AI explanations

Live demo

Purpose of the demo

- Present the XAI cardiac-risk tool in two usage modes



Modes

- **Simple:** dropdowns with predefined ranges, quick explanations
- **Expert:** free numeric inputs, advanced explanations

Demo cases

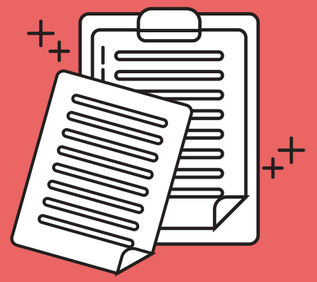
- **Low risk:** “low risk” prediction + Quick Explanation
- **High risk:** “high risk” prediction + Advanced Interpretation

Key Insights: Round 1

- Confusion over medical labels (e.g. “Sys” ≠ “Systolic”)
- Poor theme clarity & contrast (light mode hard to read)
- No loading indicator: users unsure if it was processing
- Weak patient vs. doctor separation (confusing scroll panel)
- Too many percentages, no “certainty”

Key Insights: Round 2

- Added welcome block & Basic/Advanced toggle
- Improved contrast and added parameter tooltips
- Introduced loading spinner and a “baseline” tooltip
- Redesigned Quick Explanation as bullet points



Key Insights: Results

Round 1

- Frustration with technical jargon and lack of system feedback
- Depended heavily on doctor/scientist assistance

Round 2

- All participants found the welcome page immediately
- Improved contrast & Basic/Advanced toggle praised
- Users navigated independently without help

Conclusion

Welcome block
Immediate context for users

Basic/Advanced toggle
Adaptive interface

Quick Explanation (bullets)
Quick grasp of results

Loading spinner
Instant processing feedback

Contextual tooltips
Terminology explanation

Improved contrast
Ensures readability in all themes

Thanks for the attention!

GROUP #6