

PART 1 UI — NG12 Cancer Risk Assessor

3 “Run NG12 Assessment” (Inputs)

Inputs shown

1. Patient ID

PT-110

2. Top-K

5

3. Assess button

Copy code

Input → Requirement Mapping

Patient ID

- Matches **structured data tool call**
- Retrieves patient data from `patients.json`
- Simulates BigQuery / clinical system lookup

Requirement:

“Agent calls a tool to fetch structured data (Age, Symptoms) from a simulated database”

Fulfilled

Top-K

- Controls how many NG12 chunks are retrieved from the vector store
- Allows evaluators to test:
 - Recall sensitivity
 - Retrieval confidence

Requirement:

"Agent searches a Vector Store for relevant sections of the Full NG12 PDF"

- ✓ Fulfilled
-

 **Assess Button**

Triggers:

1. Tool call → patient record
2. RAG retrieval → NG12 PDF
3. Reasoning → guideline criteria
4. Structured JSON response

Exactly matches the **logic flow described in the spec.**

 **Result Panel (Outputs)****Output Components** **Assessment Badge**

nginx

Urgent Referral

Requirement:

"Determine if the patient meets criteria for 'Urgent Referral' or 'Urgent Investigation'"

 Copy code

- ✓ Correct output classification
-

 **Confidence Score**

makefile

Confidence: 75%

- Derived from retrieval score aggregation
- Shows reasoning transparency

 Copy code

- Not hallucinated
 - ✓ Shows responsible clinical AI design
-

Reasoning Box

sql

Aged **45** and over with visible haematuria meets suspected cancer pathway re

- Concise, guideline-based
- No invented thresholds
- Matches NG12 wording

 Copy code

Requirement:

"Agent synthesizes structured data + guidelines to determine risk"

- ✓ Fulfilled
-

5 Retrieval Debug (Explainability)

Shown Data

json

```
{
  "count": 5,
  "top_score": 0.72,
  "k_score": 0.68,
  "query": "visible haematuria urology suspected cancer..."
}
```

Why this is important

- Demonstrates **observability**
- Shows this is a **real RAG system**
- Helps reviewers debug relevance

This is not explicitly required — it is a **strong bonus**

 Copy code

6 Citations Section (Critical Requirement)

Displayed

- Source: NG12 PDF
- Page number (e.g., p.59)
- Chunk ID (e.g., ng12_0059_00)
- Exact excerpt text

Requirement Mapping

"Returns a JSON response with the assessment and specific guideline citations"

- ✓ Fully satisfied
- ✓ Page + chunk-level precision
- ✓ No generic citations

This is **one of the strongest parts of your submission.**

PART 2 UI — Conversational NG12 RAG

7 Chat Settings (Inputs)

Inputs

1. Session ID

e2e-1

2. Top-K

5

3. Buttons:

- New session

 Copy code

- Load history
 - Clear session
-

Why this matters

Session ID

- Enables **multi-turn memory**
- Supports follow-ups like:
"What if the patient is under 40?"

Requirement:

"Memory must be used to support follow-ups"

Fulfilled

Load / Clear Session

- Maps to:
 - GET /chat/{session_id}/history
 - DELETE /chat/{session_id}

Exactly as requested

8 Conversation Panel (Output)

Behavior

- Shows:
 - User questions
 - Assistant answers
 - Citations per answer
- If no evidence:
 - Agent explicitly says so

Requirement:

"Refuse or qualify answers if retrieval returns insufficient evidence"

- ✓ Fulfilled (seen in guardrail tests)
-

9 Chat Input Box

Input

CSS

Ask a question about NG12...

Expected Questions

 Copy code

- "Summarize referral criteria for visible haematuria"
- "Does dysphagia require urgent referral?"
- "What age thresholds apply?"

Requirement Mapping

"Build a chat interface where a user can ask questions using NG12 content"

- ✓ Fully met
- ✓ Uses same vector DB (no re-embedding)
-

🔒 Guardrails & Grounding (Implicit but Demonstrated)

- Chat refuses chemotherapy questions
- No hallucinated treatments
- Answers always backed by citations

Requirement:

"Avoid inventing thresholds or criteria not supported by retrieved chunks"

- ✓ Strongly satisfied
-

🧠 Why This UI Scores Well

Reviewers will see: