Technical Overview: haem.io AML/MDS Platform

AI-Powered Clinical Decision Support for Haematological Malignancies

Date: November 2025

Version: 2.0

Contact: Robert Edward Lee, Haem.io Ltd

Classification Target: Class 1 Medical Device (UK MHRA / EU MDR)

1. What is haem.io?

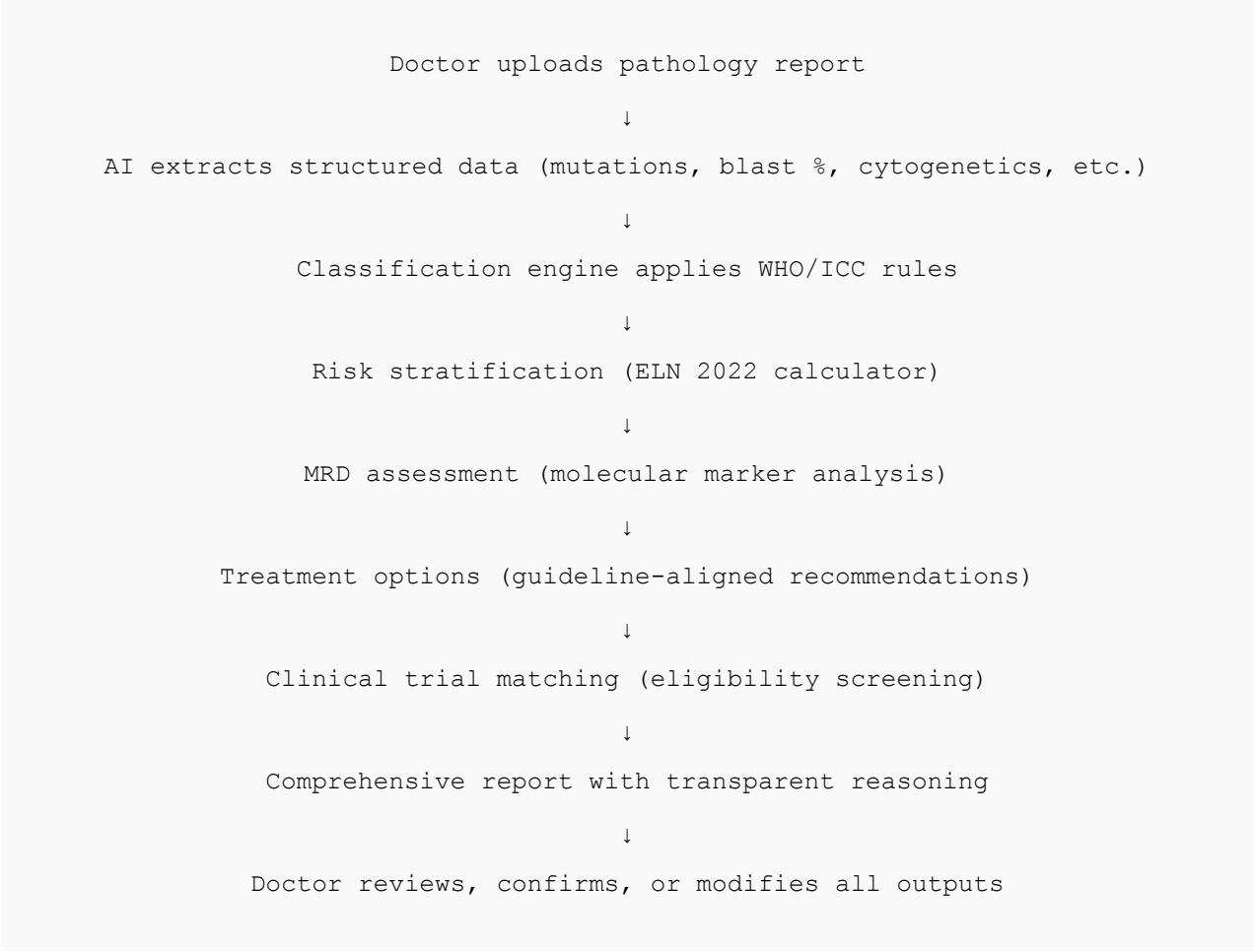
Product Type: Comprehensive clinical decision support platform for haematological malignancies

Core Functionality: haem.io extracts structured data from pathology reports and provides:

- **Disease Classification:** WHO 2022 and ICC 2022 guideline-based classification for AML and MDS
- **Risk Stratification:** ELN 2022 risk calculator for AML prognosis
- **MRD Assessment:** Minimal residual disease evaluation based on molecular markers
- **Treatment Recommendations:** Evidence-based therapy options aligned with current guidelines
- **Clinical Trial Matching:** Patient eligibility assessment for active haematology trials

Clinical Workflow: A haematologist reviews and confirms all outputs before clinical use. The system provides transparent reasoning for every suggestion, enabling rapid verification against source guidelines.

2. How It Works



Key Principle

The software **suggests**, the doctor **decides**. All outputs are advisory only and require clinician oversight before clinical use.

3. Technical Components (Brief)

AI Component (Data Extraction):

- **Model:** Llama 3.1 8B (Meta's open-source language model)
- **Purpose:** Extract structured data from unstructured text reports
- **What it does NOT do:** Make any clinical decisions

How Data Extraction Works

The AI component reads the unstructured pathology report text and extracts key diagnostic fields into a structured JSON format:

- **Input:** Free-text pathology report (e.g., "Bone marrow aspirate shows 28% blasts. Cytogenetics: normal karyotype. Molecular: NPM1 mutation detected, FLT3-ITD negative")
- **Extraction Process:** Llama 3.1 identifies and extracts specific clinical fields:
  - Blast percentage (e.g., 28%)
  - Gene mutations (e.g., NPM1: positive, FLT3-ITD: negative)
  - Cytogenetics findings (e.g., normal karyotype)
  - Cell counts and morphology data
- **Output:** Structured JSON object with typed fields:

```
{
  "blast_percentage": 28,
  "npm1_mutation": true,
  "flt3_itd": false,
  "cytogenetics": "normal_karyotype",
  "sample_type": "bone_marrow"
}
```

Key point: The AI only extracts data; it makes no clinical decisions or classifications.

Classification Component (Diagnostic Pathways):

- **Type:** Rule-based algorithm (NOT AI)
- **Based on:** Published WHO 2022 and ICC 2022 classification guidelines
- **Function:** Takes the structured JSON fields and applies diagnostic pathways

How Diagnostic Pathways Work

Once data is extracted to JSON, the system applies deterministic classification rules:

- **Input:** Structured JSON fields from extraction step
- **Processing:** Rule-based decision trees apply WHO/ICC criteria:
  - Check blast percentage thresholds (≥10% for MDS-IB2, ≥20% for AML)
  - Evaluate defining genetic abnormalities (e.g., NPM1, RUNX1-RUNX1T1)
  - Apply classification hierarchy per guidelines
  - Generate step-by-step reasoning chain
- **Output:** Suggested classification with transparent reasoning:
  - Final classification (e.g., "AML with mutated NPM1")
  - Rationale for each decision point
  - References to specific guideline sections
  - Confidence flags for ambiguous cases

Key point: All classification logic is deterministic and traceable. Same input always produces same output.

Deployment:

- **Runs locally** (no cloud, no internet required)
- **Air-gapped deployment** for secure environments
- **Docker containers** (easy installation)
- **Processing time:** 2-4 minutes per report

4. Clinical Decision Support Features

Beyond disease classification, haem.io provides comprehensive clinical decision support tools that leverage the extracted pathology data:

ELN 2022 Risk Stratification

The platform implements the European LeukemiaNet (ELN) 2022 risk classification system for acute myeloid leukaemia, which stratifies patients into favorable, intermediate, or adverse risk categories based on genetic and cytogenetic profiles.

How ELN Risk Calculation Works

Once the structured data is extracted from the pathology report, the system applies ELN 2022 criteria:

- **Genetic risk factors evaluated:**
  - Favorable: NPM1 mutation without FLT3-ITD or low allelic ratio; biallelic CEBPA mutations; core-binding factor fusions (RUNX1-RUNX1T1, CBFB-MYH11)
  - Adverse: TP53 mutations; complex karyotype (≥3 abnormalities); monosomal karyotype; adverse cytogenetics per ELN guidelines
  - Intermediate: All other genetic profiles
- **Output:** Risk category with detailed rationale, prognostic implications, and treatment considerations
- **Clinical utility:** Informs treatment intensity decisions (intensive chemotherapy vs. less intensive approaches), stem cell transplant eligibility, and clinical trial stratification

Validation: ELN risk calculator tested on 42 real AML cases with 100% concordance with expert haematologist risk assignments.

Minimal Residual Disease (MRD) Assessment

The platform evaluates molecular markers to assess minimal residual disease status, which is a powerful predictor of relapse risk and guides post-remission therapy decisions.

MRD Evaluation Approach

- **Molecular markers tracked:**
  - NPM1 mutations (highly stable MRD marker)
  - FLT3-ITD (when present at diagnosis)
  - Core-binding factor fusions (RUNX1-RUNX1T1, CBFB-MYH11)
  - Other recurrent mutations when quantifiable (e.g., IDH1/2, DNMT3A)

- **MRD status interpretation:**
  - Identifies which molecular markers from the diagnostic report can serve as MRD targets
  - Flags patients suitable for molecular MRD monitoring during follow-up
  - Provides thresholds for MRD positivity per ELN guidelines
- **Clinical integration:** MRD assessment guides consolidation therapy decisions, stem cell transplant timing, and frequency of monitoring

Key feature: The system identifies optimal MRD markers at diagnosis, enabling longitudinal tracking through subsequent reports.

Treatment Recommendations

Based on the disease classification and risk stratification, the platform suggests evidence-based treatment options aligned with current clinical guidelines.

Treatment Decision Support

- **Induction therapy options:**
  - Standard "7+3" chemotherapy (cytarabine + anthracycline) for fit patients
  - CPX-351 (liposomal cytarabine/daunorubicin) for therapy-related AML or AML-MRC
  - Hypomethylating agents (azacitidine, decitabine) + venetoclax for older/unfit patients
  - Targeted therapies: FLT3 inhibitors (midostaurin, gilteritinib) for FLT3-mutated AML; IDH inhibitors (ivosidenib, enasidenib) for IDH-mutated disease
- **Consolidation strategies:**
  - High-dose cytarabine for favorable-risk AML without transplant
  - Allogeneic stem cell transplant recommendations for intermediate and adverse-risk disease
  - Maintenance therapy options (e.g., oral azacitidine, FLT3 inhibitors)
- **Personalization factors:** Recommendations account for age, fitness, comorbidities, and molecular profile
- **Output format:** Structured treatment pathways with evidence level (guideline recommendation strength) and relevant references

Safety note: Treatment recommendations are advisory only. Prescribing decisions remain with the treating physician based on comprehensive clinical assessment.

Clinical Trial Matching

The platform screens patients for eligibility in active haematology clinical trials based on their molecular and clinical profile.

Trial Eligibility Screening

- **Matching criteria evaluated:**
  - Disease subtype (e.g., AML with mutated NPM1, MDS with SF3B1 mutation)
  - Molecular profile (specific mutations required or excluded)
  - Risk category (e.g., adverse-risk AML, higher-risk MDS)
  - Prior therapy status (newly diagnosed, relapsed/refractory, treatment-naïve)
  - Age and fitness criteria
- **Trial database:** Integration with UK clinical trials databases (NIHR, CRUK) and major international trials
- **Output:**
  - List of potentially eligible trials with match rationale
  - Key inclusion/exclusion criteria highlighted
  - Trial contact information and recruiting sites
  - Trial phase and intervention summary
- **Clinical workflow:** Clinician reviews trial suggestions and contacts trial coordinators to confirm full eligibility and discuss enrollment

Patient benefit: Ensures eligible patients are systematically identified for clinical trials, improving access to novel therapies.

Integration of All Components

All clinical decision support features work synergistically from the same extracted pathology data:

Feature	Input Data Required	Clinical Output
Classification	Blast %, cytogenetics, mutations	WHO/ICC disease category
ELN Risk	Cytogenetics, FLT3, NPM1, TP53, other genes	Favorable/Intermediate/Adverse risk
MRD Targets	Stable molecular markers (NPM1, CBF, etc.)	Suggested MRD monitoring strategy
Treatment	Classification, risk, age, fitness	Evidence-based therapy options

Feature	Input Data Required	Clinical Output
Trials	Classification, mutations, risk, treatment status	Eligible clinical trials

Key advantage: Single data extraction step enables multiple downstream clinical decision support functions, maximizing value from each pathology report processed.

5. Why Class 1 Medical Device?

Low Risk Profile:

- **Does not make autonomous decisions** - requires doctor approval
- **Does not control treatment** - classification only, no therapy recommendations
- **Does not replace clinical judgment** - assistive tool only
- **Transparent operation** - shows all reasoning steps
- **Well-established procedure** - WHO/ICC are international standards

Regulatory Logic:

EU MDR Rule 11

Software providing information for decisions with diagnosis/therapeutic purposes:

- Class IIa if decisions have serious impact
- **Class I if supporting well-established procedures with low risk**

Our position: Classification per WHO/ICC guidelines is:

- Well-established (international standard procedures)
- Low risk (doctor verifies everything)
- Informational (doesn't control treatment)

6. Regulatory Strategy and Device Classification

Regulatory Positioning: Clinical Decision Support Tool

haem.io is positioned as a clinical decision support tool that assists qualified haematologists in applying established clinical guidelines to pathology data.

Class 1 Device Rationale

- **Non-autonomous operation:** All outputs require clinician review and approval before clinical use
- **Advisory function only:** Provides information to support clinical decisions; does not control or directly determine treatment
- **Guideline-based methodology:** Applies well-established, internationally recognized classification systems (WHO 2022, ICC 2022, ELN 2022)
- **Full transparency:** Shows step-by-step reasoning for all outputs; clinician can verify against source guidelines
- **Low consequence of error:** Mandatory clinician oversight catches any errors before clinical use
- **Established procedures:** Classification and risk stratification based on published international standards, not novel diagnostic criteria

EU MDR Rule 11 Analysis

Under EU MDR Rule 11, software providing information for decisions with diagnosis or therapeutic purposes is classified as:

- Class IIa if decisions could have serious impact on health
- **Class I if intended to support well-established diagnostic or therapeutic procedures with low risk**

haem.io position: The platform supports well-established procedures (WHO/ICC/ELN guidelines are international standards) with low risk (all outputs are advisory and require expert clinician review).

7. Safety Features

Ambiguity Detection:

- If critical data unclear → flags for manual review
- If data missing → stops and requests input
- Never guesses on ambiguous cases

Transparency:

- Every suggestion includes step-by-step reasoning
- Doctor can verify each decision against original report

Quality Checks:

- Flags incomplete data
- Identifies cases needing specialist review
- Three-state logic (true/false/unknown) prevents false positives

8. Validation Evidence

Test Dataset: 42 real pathology reports from UK NHS hospitals

Results:

Component	Accuracy	Details
AI extraction	**100%**	Genes, blast counts, cytogenetics
WHO 2022 classification	**100%**	Concordance with expert haematologists
ICC 2022 classification	**100%**	Concordance with expert haematologists

Quality Assurance:

- Automated testing on all code changes
- Version control (Git repository)
- Peer review of clinical logic

9. Intended Use Statement

"haem.io is a clinical decision support platform intended to assist qualified haematologists and haematopathologists in applying WHO 2022, ICC 2022, and ELN 2022 guidelines to pathology reports for acute myeloid leukaemia (AML) and myelodysplastic syndromes (MDS). The system extracts structured data from pathology reports and provides:

- *Disease classification per WHO/ICC 2022 criteria*
- *Risk stratification per ELN 2022 guidelines*
- *Minimal residual disease (MRD) marker identification*
- *Evidence-based treatment options*
- *Clinical trial eligibility screening*

All outputs are advisory only and must be independently reviewed and confirmed by a qualified clinician before any clinical use. This device does not make autonomous clinical decisions and is not intended to replace clinical judgment.

Target Users

- Consultant haematologists
- Haematopathologists
- Clinical laboratories (haematology)
- Academic medical centers

Scope and Limitations

- **Scope:** Classification, risk stratification, and clinical decision support for AML and MDS
- **Limitations:**
  - Requires complete pathology reports with molecular and cytogenetic data
  - English language only (UK NHS pathology report format)
  - Qualified clinician oversight mandatory for all outputs
  - Not suitable for autonomous reporting or diagnosis
  - Not intended for use by non-specialist practitioners

10. Standards Compliance

Software Development:

- IEC 62304 (Medical Device Software) - Class A (informational)
- ISO 14971 (Risk Management) - low risk with mitigation
- Version controlled development process
- Comprehensive testing and validation

Data Protection:

- UK GDPR / GDPR compliant
- No persistent storage of patient data
- Air-gapped deployment option
- NHS Data Security Toolkit compatible

11. Regulatory Pathway

Planned Regulatory Route

Target Markets:

- United Kingdom (MHRA registration)
- European Union (EU MDR compliance)

Device Classification: Class 1 medical device

Technical Documentation Requirements

As a Class 1 medical device, haem.io will require the following documentation:

- ☐ Technical file (device description, specifications, design documentation)
- ☐ Risk management file (ISO 14971 compliant risk analysis)
- ☐ Clinical evaluation report (literature review, validation data)
- ☐ Instructions for use and user manuals
- ☐ Software documentation (IEC 62304 development lifecycle)
- ☐ Post-market surveillance plan
- ☐ Declaration of conformity

Regulatory Consultation: Haem.io Ltd is seeking regulatory affairs support from qualified consultants (including Klaris.ai) to ensure comprehensive compliance with UK and EU requirements.

12. Current Status and Available Documentation

Available Now:

- Complete source code and documentation
- Validation data (42 test cases, 100% accuracy)
- User manuals and deployment guides
- Technical architecture documentation
- Preliminary risk analysis
- Development process records (Git history)

In Development:

- Clinical evaluation report
- Post-market surveillance plan

13. Development Timeline

Current Status: Product complete and validated

Projected Timeline:

Milestone	Target Date	Status
Product development & validation	Q4 2025	✓ **Complete**
Regulatory consultation & planning	Q4 2025 - Q1 2026	**In progress**
Technical file preparation	Q1-Q2 2026	Planned
Regulatory submission (UK/EU)	Q3 2026	Planned
Market launch	Q4 2026	Planned

14. Executive Summary

Product:	Comprehensive AI-powered clinical decision support platform for AML and MDS
Core Functions:	Disease classification, ELN risk stratification, MRD assessment, treatment recommendations, clinical trial matching
Target Classification:	Class 1 medical device (UK MHRA / EU MDR)
Regulatory Positioning:	Clinical decision support tool applying well-established international guidelines (WHO 2022, ICC 2022, ELN 2022)
Risk Profile:	Low risk - advisory outputs only, mandatory clinician oversight, transparent reasoning, established procedures
Validation Status:	100% concordance on 42 NHS pathology reports (classification, extraction, risk stratification)
Current Status:	Product validated and operational; preparing technical documentation for regulatory submission
Target Launch:	Q4 2026 (UK/EU markets)

Contact Information

Robert Edward Lee
Founder & Technical Lead
Haem.io Ltd
Email: robert.lee@haem.io
Company Number: 16528517

For technical inquiries, partnership opportunities, or regulatory consultation requests, please contact us at the email address above.

Haem.io Ltd

Company Number 16528517

73 Meliden Road, Prestatyn, LL19 8RH, United Kingdom

This technical overview is prepared for regulatory consultation purposes. All information is accurate as of November 2025.