

G-Cloud 15

Service Definition

hepatoSIGHT.



January 2026

The problem that hepatoSIGHT™ helps to address

Liver disease (which has many causes) places a huge burden on global public health. According to the British Liver Trust, one in nine people in the UK suffers from the condition, and it causes more than 10,000 deaths each year. Most alarmingly, there has been a four-fold increase in death rates over the last 50 years, when most other death rates like heart disease, cancer, and diabetes have been in decline.

There are three key facts about the liver that most people are not aware of:

- It is usually silent: Liver disease usually doesn't show symptoms until severe damage has already happened, making early diagnosis and treatment challenging. Patients frequently present with advanced and incurable disease, yet the signs of early liver disease are often hidden in previous routine test results.
- It is common: Liver disease is a major cause of illness and premature death. It affects around one in nine people, including many people who don't consume alcohol. Liver disease is caused by a range of factors spanning metabolic syndrome and obesity, viral infection, autoimmune conditions, and substance abuse. In addition, liver disease (and the growth in cardiometabolic disorders) is now widely recognised as one of the greatest global public health problems, with one of the biggest issues being its late diagnosis.
- It is treatable: The liver has an amazing ability to heal, it can even grow back whole sections after surgery. However, if liver disease is too advanced, the scarring stops the regeneration process. The key is finding problems early and getting people on treatments as soon as possible. Targeted lifestyle changes are often highly effective at improving liver health.

These three facts combine to create a perfect storm of liver disease in the population, and set the scene for the idea behind Predictive Health Intelligence™.

The clinical problem that is frequently seen and needs solving is quite clear:

- The patient perspective: they often do not understand why they are found so late in their disease's progression. This is often highly distressing.
- The clinician perspective: they are frustrated at seeing people so late when treatment options are far more limited, despite there being signals that advanced disease is likely.

After decades of seeing this familiar, heart-breaking story unfold, there was a tipping point for Dr Tim Jobson. One day in early 2018 he saw a newly referred patient in his liver clinic at Somerset NHS FT. The patient was the same age as Dr Jobson and had never been seen by a liver doctor before. When Dr Jobson looked back at his records, through the lens of hindsight, it was clear that for several decades his liver was becoming more and more injured and that disaster was imminent, yet no one had put the pieces together in time. On their first interaction Dr Jobson diagnosed inoperable, incurable liver cancer.

This case, whilst by no means unique, stayed with Dr Jobson and compelled him to find a way to reach these same sorts of patients, today. Not in 5 or 10 years when there is nothing anyone can do about their condition. If Dr Jobson could look back through a person's medical records and see the train coming down track, why couldn't we build a system to do the same thing at scale?

With this germ of an idea, Dr Jobson worked with colleagues at Predictive Health Intelligence to develop hepatoSIGHT – an intuitive browser that enables clinicians to easily identify people for whom certain patterns of blood test results are present.

hepatoSIGHT is a powerful, easy to use tool that enables clinicians to identify cohorts of patients with user-defined patterns of blood test abnormalities. It is a highly flexible tool that allows the user to specify parameters indicative of potential liver disease and set thresholds for these parameters to flexibly risk stratify the population.

This tool allows rapid, user defined searches of a large database of pre-existing blood tests (currently 600,000 individuals, 205 million tests). The system is designed so that it can be expanded; ultimately to cover the whole of the country's population.

hepatoSIGHT includes data to target deprivation, and to accurately map cohort sizes before they are generated. These features allow targeted evaluation and intervention within specific populations.

The system can review 100s of millions of lines of data within seconds based on patterns of abnormalities and thresholds set by the healthcare team, finding patients who meet specified criteria. It is an intuitive, easy-to-use system that rapidly risk stratifies the population and identifies individuals who will benefit from intervention. It is also an effective tool for identifying cohorts of patients for potential enrolment in research studies. This system is secure and allows the frontline healthcare team to interrogate hundreds of thousands of patient records to create manageable cohorts for intervention in seconds.

hepatoSIGHT supports the identification of for longitudinal trends over time – e.g. trends, integrals, maxima and minima etc. This approach has been shown to be extremely powerful, with an odds ratio for the detection of liver disease > 18 in preliminary tests (Wesley E, Matull W, Kitchin A, et al - A cumulative liver damage index (CLDI) identifies patients at risk of significant liver disease Gut 2022;71:A4).

hepatoSIGHT has been in use in Somerset since early 2023 and is routinely used to find people most at risk of liver disease. In the first year, over 700 people have been contacted and invited to make an appointment in the liver clinic for further investigation, none of whom were already under the care of a liver specialist.

Service Specification

Introduction

- This service specification outlines the requirements for the provision of a bespoke hepatoSIGHT IT case-finding system. This service is not a supply of off-the-shelf software or staff. It is a complete, managed IT case-finding service.
- The service includes the design, development, implementation, and support of a tailored IT case-finding solution to meet the specific requirements of the NHS.

Service Description

- The service is the provision and ongoing support of a complete IT case-finding system designed to the specifications of the NHS. This system is designed to enable clinicians to identify cohorts of patients with user-defined patterns of blood test abnormalities, specifically for liver disease.
- The core component of the service is the hepatoSIGHT platform, a bespoke software solution tailored for identifying patients with potential liver disease. This platform is not a pre-existing product as extensive data management and configuration is required to enable it to function.
- The service includes providing the necessary computer infrastructure, with data hosted in HSCN compliant cloud servers, managed by the supplier.
- The service also covers the development, configuration, delivery, and ongoing support of the bespoke hepatoSIGHT software.

Software Development and Configuration

- The service includes the initial build and ongoing refinement of the hepatoSIGHT platform. This is software that is substantially modified to meet the specific needs of the NHS, with modifications made for each implementation.
- The software development includes configuration of the system to receive data from multiple sources.
- The service involves mapping test IDs to the specific results required and configuring the system based on the supplied location mapping.
- Ongoing support includes providing updates, helpdesk facilities, and troubleshooting any functionality issues.
- The system is designed to review millions of lines of data within seconds, based on parameters set by the healthcare team.
- This includes support of bespoke software which is part of the overall IT case-finding system

- The system can search for longitudinal trends over time as defined by the requirements of the NHS users.

Data Acquisition and Management

- The service encompasses all aspects of data acquisition, mapping, and configuration.
- The service includes the determination of the most suitable data source for each site.
- The service ensures data is acquired for the agreed geographic footprint, usually an Integrated Care System (ICS).
- Data can be taken from multiple sources, but the minimum number possible will be used.
- Historical data will be imported as far back as practicable.
- The service manages the test ID process, mapping the test IDs to the required test results, and then enhancing the dataset with bespoke calculations to create composite test results, and patient level calculations.
- The process involves extracting data from source systems to a secure, HSCN-compliant staging server before importing it into the hepatoSIGHT system.
- The system will be configured based on numerous features including test location codes and bespoke classifications of non-numeric results.

IT Infrastructure

- The hosting of the service will be on a secure, HSCN-compliant cloud-based infrastructure.
- This includes all necessary migration of data from the recipient to the cloud-hosted service.

Project Management and Support

- The service includes project management to ensure each implementation is successful.
- The service provides ongoing software service and support for a period of 3 years.
- This involves ongoing support to the local team, addressing any data or functionality issues.

- The service includes support with local governance processes, including Data Protection Impact Assessments (DPIA), Data Sharing Agreements (DSA's) and Confidentiality Advisory Group (CAG) submissions).
- The service includes support in obtaining necessary permissions and managing governance issues.

Staffing

- Any staff provided will be under the direction and control of the supplier.
- Staff are an integral part of the provision of the IT case-finding services, not a separate supply of staff.
- The staff's activities are a necessary part of the overall provision of the IT case-finding system.

Exclusions

- The service specifically excludes the supply and support of off-the-shelf software which may be required to use hepatoSIGHT (e.g. Microsoft Office products, web browsers).
- The service excludes the hire of computer consultants to add expertise to in-house teams where those consultants are managed by the in-house teams.
- This service does not include web design services.

Contractual Structure

- All components of this contract are part of the provision of a single IT case-finding system.
- Multiple suppliers are involved. These are managed under one contract to provide an IT case-finding system to the specifications of the NHS.
- All services provided are integral to the overall IT case-finding system procurement.

Governance and Compliance

- The system will be configured so that only clinicians with legitimate rights to patient identifiable data within their locality can access the information.
- Patient identifiable data is only imported once a local DPIA is complete and approved, and DSAs signed for each Data Controller.
- The service adheres to the highest levels of Information Governance and Information Security.
- The sub-contractors are fully compliant with NHS security and IG requirements.

- The system has undergone extensive testing.
- The system has been evaluated in line with the appropriate clinical safety standard (DCB0129).
- The platform is certified as a MHRA Class 1 device.

Implementation approach

Predictive Health Intelligence will work with customers to identify the best data source for blood test results – which is typically within the Laboratory Information Management Systems (LIMS). Part of the implementation planning will be to agree how much historic data will be loaded into hepatoSIGHT.

Steps in the implementation process are described below.

Data acquisition

The process for bringing data on board may vary for each site depending on the nature of the project; this section provides a high-level summary of the approach and should be used to develop the detailed local project plan for a new site.

Data source

For each new site there should be a determination of the most suitable source for the project data. This should take into account the data items required (Appendix B) and the fields required (Appendix C). The aim is to acquire data for the agreed geographic footprint (usually an ICS). The system can take data from multiple, different sources for a single footprint, but the minimum number possible should be used. The aim should also be to import as much historical data as possible (i.e. going back in time as far as is practicable). Again, this may require imports from multiple sources with a balance being struck between the number of sources and the earliest historic data to be imported.

Local DPIA and Governance

Once agreements to proceed are in place, a local DPIA will be produced. Appropriate permissions will also be sought if there is a research component (usually via the IRAS process).

For more information on governance for hepatoSIGHT, see Appendix 1.

Test ID process

Prior to the transfer of any data, a complete list, or catalogue, of test IDs from each data source is required. Depending on the local test record system in use, a single biochemical/pathological result may have multiple test IDs. Providing a comprehensive

list is produced, this matter is managed by mapping data items within the hepatoSIGHT system.

The complete list of test IDs will be analysed by the PHI team and mapped to the test results required (typically as per Appendix B). Once this is complete, the required test ID list will be generated. This defines the data items required and will be passed back to the host site to enable an appropriate extract of data.

Units and normal ranges

From the list of required test IDs the host site will generate a list of units (required) and normal ranges (where possible, but not essential) mapped to each test ID.

This list will be used by the PHI team to configure the hepatoSIGHT database in preparation for the data import process.

Trial data extract

From each data source, a trial data extract for the specified test IDs should be obtained and passed to the project team. This can be relatively small (e.g. 1 week of data for each source system).

This data set should be irreversibly anonymised (i.e. NHS number not visible). If possible, this should be supplied as 'linked anonymised' data – i.e. with an irreversible transformation of the NHS number, although this is not essential.

The trial extract will be analysed by the PHI team to determine whether there are likely to be any additional complexities for the subsequent full import of data. An initial review for data completeness will be made at this stage to ensure the viability of the data source(s) for the project.

Full data extract

Once all relevant permissions for analysis of patient identifiable data are in place and the local DPIA has been approved, a full extract of data from the source system(s) will be performed. This will be an extract of all results for the agreed list of test IDs.

For previous implementations this has been approximately 140m items over 15 years of data for a population of 0.5m.

This data extract will be passed to the secure, HSCN-compliant staging server, in preparation for import into the hepatoSIGHT system.

Test location codes

A complete list of test location codes will be extracted by the PHI team from the data on the staging server. These will need to be mapped to the following locations:

- Primary Care
- Secondary Care – outpatient
- Secondary Care – inpatient
- Other
- Unknown

The PHI team will work with the local team to complete the mapping process. The hepatoSIGHT system will then be configured as part of the data import process, based on the supplied location mapping.

Non-numeric results

A complete set of non-numeric results for each test will be extracted by the PHI team from the data on the staging server.

This list will be analysed by the PHI clinical team. In general these items are standard clinical terms or self-explanatory (e.g. with unusable results) but any non-standard items that cannot be identified will be discussed further with the local team for clarification.

In general these are mapped as:

- 'usable' vs 'unusable'
- 'positive [abnormal]' vs 'negative [normal]'

The hepatoSIGHT system will then be configured, as part of the data import process, on the basis of this mapping of non-numeric results.

Data import and configuration

Once the above steps are complete, all data importing, mapping and configuration will be undertaken by the PHI team.

Testing

Testing will be agreed, in advance, as part of the project. This will normally involve searches based on a subset of key algorithms to ensure appropriate results are generated - along with a small number of patient results directly compared to live clinical data.

Full regression testing will not normally be undertaken at this stage, as the system has already been fully tested.

Appendix 1 - Governance

The Predictive Health Intelligence team has worked to ensure the highest levels of Information Governance and Information Security have been applied to project at all stages. These are summarised in this section and more detail can be supplied as needed.

Software rights

The software was developed as part of a National Institute for Health and Care Research (NIHR) funded project (i4i – NIHR200965). All intellectual property rights are owned within the public sector (Somerset NHS FT) and provided under licence to Predictive health Intelligence for provision of services.

Data access

Predictive Health Intelligence Limited designed the software and manages the project but has no access to patient identifiable data. The system is configured to ensure that only clinicians with legitimate rights to patient identifiable data within their locality can access this information.

Patient identifiable data will only be imported into the system once a local DPIA is complete and approved.

Sub-contractors

Predictive Health Intelligence works with highly specialised sub-contractors which are fully compliant with NHS security and IG requirements.

Stalis Ltd

Stalis Ltd developed the hepatoSIGHT search tool. This company has extensive experience with migrating and analysing large data sets within the NHS. The company has the following accreditations:

- Approved supplier working with HSCN
- ISO 27001
- Cyber Essentials Plus (CE+) certification

ARO Ltd

ARO Ltd is the supplier of the cloud hosting solution for hepatoSIGHT.

This company is fully accredited to manage patient identifiable data in a cloud hosting solution within HSCN.

Testing

The hepatoSIGHT solution has undergone extensive testing by the supplier and 'User Acceptance Testing' by the clinical team at Somerset NHS Foundation Trust.

Results of testing are available for inspection.

Clinical safety evaluation

The system is not designed to support individual clinical decision making. It is a tool to identify cohorts of patients potentially at risk, in a highly sophisticated way with little or no technical data expertise required by the end user.

Once potential patients are identified with the system, all clinical interventions should be based on standard clinical evaluation of the patient (history, examination etc) along with data existing within local clinical systems.

This approach ensures that there is very limited risk to clinical safety. Nonetheless the system has been evaluated in line with the appropriate clinical safety standard (DCB0129) and a Clinical Safety Case Report is available for review.

Medical Device Regulations

The system is a database that is optimised for identifying and risk stratifying cohorts of patients based on existing blood test results.

It is never used to support individual clinical decision making. It does not generate or indicate any specific diagnoses, nor does it recommend any treatments.

hepatoSIGHT is MHRA certified as a Class I Medical Device.

Appendix B – Tests required

The following table describes the tests for which results are required for a typical hepatoSIGHT implementation. Test codes and test names will vary according to data source, and all mapping will be completed within the hepatoSIGHT system.

NAME OF FIELD
Name
DOB
NHS NUMBER
POST CODE
DATE OF TEST
SOURCE CODE
BILIRUBIN
ALBUMIN
AST
ALT
GAMMA GT
ALKALINE PHOSPHATASE
HB
PLTS
MCV
RDW
WBC
RBC
HCT
MCH
MCHC
Neutrophils

Lymphocytes
Monocytes
Eosinophils
Basophils
LUC
PT
INR
APTT
APTT RATIO
AFP
CA 199
CA 125
CA 153
CEA
FERRITIN
IRON
COPPER
24hr COPPER OUTPUT
CAERULOPLASMIN
ALPHA 1 ANTITRYPSIN
A1AT PHENOTYPE
TRANSFERRIN
TRANSFERRIN SATURATION
IGG
IGA
IGM
Antinuclear Antibody
Smooth Muscle Antibodies

Mitochondrial Antibodies
ANCA SCREEN
LKM ANTIBODIES
IgA tTG
ANCA TITRE
SODIUM
POTASSIUM
UREA
CREATININE
HbA1C
GLUCOSE
CHOLESTEROL
TRIGLYCERIDES
HDL CHOLESTEROL
LDL CHOLESTEROL
MAGNESIUM
PHOSPHATE
C REACTIVE PROTEIN
CALCIUM
Endomysial Antibodies
Hepatitis A confirmation
Hepatitis A IgG
Hepatitis A IgG Antibody
Hepatitis A IgG ELISA
Hepatitis A IgM
Hepatitis A IgM Antibody
Hepatitis B confirmation tests
Hepatitis B core antibody

Hepatitis B Serology
Hepatitis B surface antigen
Hepatitis C Genotype
Hepatitis C PCR
Hepatitis C Serology
Hepatitis E PCR (Faeces)
Hepatitis E Serology
Hep B core IgM (Vitros)
Hep C RNA (qualitative PCR)
HEP2 ANA

Appendix C – fields required

The following table describes the fields required for each of the test items (Appendix 3).

The hepatoSIGHT initial data import will be a table of results with one row per numeric or text results with the fields described below. All data mapping, cleansing and aggregation will then take place within the hepatoSIGHT system.

Content of each row	TestResultID NHSNumber DOB DOD Sex Age TestID TestCode/Name DateOfReceipt DateEntered NumericResult ResultNotLarge [TEXT RESULT] SourceLocationCode SourceLocationCode_Name	This field list is the list generated from LabCentre in the original hepatoSIGHT build. Field names may vary dependent on the data source Deviation from this may require mapping, dependent on data source. Setting up the mapping should be a one-time exercise for each new data source only – and should proceed automatically for additional DSAs / time periods for a given data source
---------------------	---	---