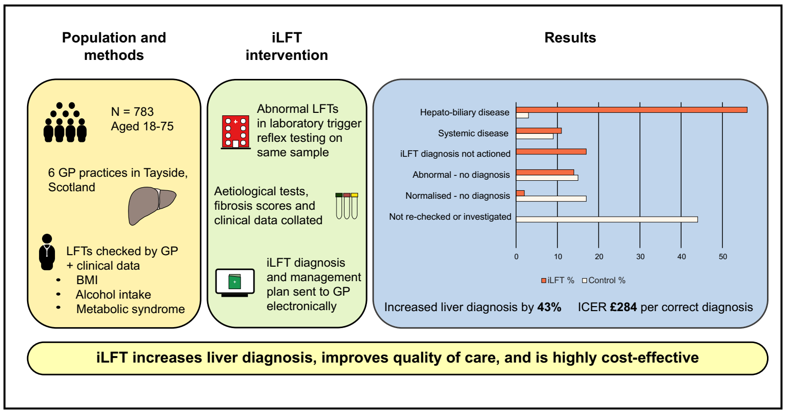
Research Article
Cirrhosis

JOURNAL
OF HEPATOLOGY

Intelligent liver function testing (iLFT): A trial of automated diagnosis and staging of liver disease in primary care

Graphical abstract



Highlights

- Intelligent liver function testing utilises the smarter application of existing knowledge and technology.
- Intelligent liver function testing increases diagnosis of liver disease by 43%, with diagnostic accuracy over 90%.
- Intelligent liver function testing enables earlier identification of treatable liver disease.

Authors

John F. Dillon, Michael H. Miller, Emma M. Robinson, ..., Peter T. Donnan, Kathleen A. Boyd, Ellie Dow

Correspondence

e.robinson7@nhs.net
(E.M. Robinson)

Lay summary

There is a growing epidemic of advanced liver disease, this could be offset by early detection and management. Checking liver blood tests (LFTs) should be an opportunity to diagnose liver problems, but abnormal results are often incompletely investigated. In this study we were able to substantially increase the diagnostic yield of the abnormal LFTs using the automated intelligent LFT system. With the addition of referral recommendations and management plans, this strategy provides optimum investigation and management of LFTs and is cost saving to the NHS.

http://dx.doi.org/10.1016/j.jhep.2019.05.033

Research Article
Cirrhosis

JOURNAL OF HEPATOLOGY

Intelligent liver function testing (iLFT): A trial of automated diagnosis and staging of liver disease in primary care

John F. Dillon[1], Michael H. Miller[1], Emma M. Robinson[1,*], Adrian Hapca[2], Mohsen Rezaeihemami[3], Christopher Weatherburn[4], Paul G. McIntyre[5], Bill Bartlett[6], Peter T. Donnan[2], Kathleen A. Boyd[3], Ellie Dow[6]

[1]*Division of Clinical and Molecular Medicine, University of Dundee, Ninewells Hospital and Medical School, Dundee, UK;* [2]*Dundee Epidemiology and Biostatistics Unit, University of Dundee, Dundee, UK;* [3]*Institute of Health and Wellbeing, University of Glasgow, Glasgow, UK;* [4]*Dundee Health and Social Care Partnership, Dundee, UK;* [5]*Department of Microbiology and Virology, Ninewells Hospital and Medical School, Dundee, UK;* [6]*Department of Clinical Sciences, Ninewells Hospital and Medical School, Dundee, UK*

See Editorial, pages 651–653

Background & Aims: Liver function tests (LFTs) are frequently requested blood tests which may indicate liver disease. LFTs are commonly abnormal, the causes of which can be complex and are frequently under investigated. This can lead to missed opportunities to diagnose and treat liver disease at an early stage. We developed an automated investigation algorithm, intelligent liver function testing (iLFT), with the aim of increasing the early diagnosis of liver disease in a cost-effective manner.

Methods: We developed an automated system that further investigated abnormal LFTs on initial testing samples to generate a probable diagnosis and management plan. We integrated this automated investigation algorithm into the laboratory management system, based on minimal diagnostic criteria, liver fibrosis estimation, and reflex testing for causes of liver disease. This algorithm then generated a diagnosis and/or management plan. A stepped-wedged trial design was utilised to compare LFT outcomes in general practices in the 6 months before and after introduction of the iLFT system. Diagnostic outcomes were collated and compared.

Results: Of eligible patients with abnormal LFTs, 490 were recruited to the control group and 64 were recruited to the intervention group. The primary diagnostic outcome was based on the general practitioner diagnosis, which agreed with the iLFT diagnosis in 67% of cases. In the iLFT group, the diagnosis of liver disease was increased by 43%. Additionally, there were significant increases in the rates of GP visits after diagnosis and the number of referrals to secondary care in the iLFT group. iLFT was cost-effective with a low initial incremental cost-effectiveness ratio of £284 per correct diagnosis, and a saving to the NHS of £3,216 per patient lifetime.

Conclusions: iLFT increases liver disease diagnoses, improves quality of care, and is highly cost-effective. This can be achieved with minor changes to working practices and exploitation of functionality existing within modern laboratory diagnostics systems.

Lay summary: There is a growing epidemic of advanced liver disease, this could be offset by early detection and management. Checking liver blood tests (LFTs) should be an opportunity to diagnose liver problems, but abnormal results are often incompletely investigated. In this study we were able to substantially increase the diagnostic yield of the abnormal LFTs using the automated intelligent LFT system. With the addition of referral recommendations and management plans, this strategy provides optimum investigation and management of LFTs and is cost saving to the NHS.

Introduction

There has been an exponential increase in the number of liver function tests (LFTs) requested in general practice.[1,2] A proportion are checked for the investigation of liver disease but most are for investigating undifferentiated illness, or monitoring non-hepatic long-term health conditions.[3–5] It is unknown how significant a solitary abnormal LFT result is;[1] does it signify current or future liver disease, disease in other organs, or is it a temporary phenomenon of little clinical relevance?[6]

Studies have shown that approximately 20% of initial LFTs are abnormal.[1] The cause of LFT abnormalities vary geographically and the cost effectiveness of additional testing will vary also. Additional diagnostic tests such as ultrasound or screening blood tests may still leave many abnormal LFTs unexplained.[2,7] Further investigations such as liver biopsy are invasive and expensive.[8] Primary care guidelines on the evaluation of abnormal liver enzyme results, in asymptomatic individuals, do not take the costs to the patient or the health service into account.[8,9]

The commonest causes of abnormal LFTs leading to chronic liver disease in the UK and most developed countries are non-alcoholic fatty liver disease (NAFLD),[10] alcohol-related liver disease (ALD),[11] and hepatitis C virus (HCV) infection.[12] All pose a

Keywords: Abnormal LFTs; Automated testing; Intelligent liver function testing; Non-invasive diagnosis; Blood tests; Liver disease.
Received 4 September 2018; received in revised form 10 May 2019; accepted 18 May 2019; available online 19 June 2019
* Corresponding author at: University of Dundee, Ninewells Hospital and Medical school, Dundee DD1 9SY, UK. Tel.: +44 01382 388685.
E-mail address: e.robinson7@nhs.net (E.M. Robinson).

considerable economic burden on the health service.[13,14] However there are many other causes of abnormal LFTs that need to be considered; including biliary disease, drug reactions, systemic illness, malignancy and hepatic infections.[15,16] Despite the increasing use of LFTs, patients continue to present with undiagnosed end-stage liver disease, which may have been preventable by earlier diagnosis.[6] Early detection may improve prognosis and treatment options, whilst diagnostic delay may be damaging to patients and to professional reputations.

There have been few population studies to date that have quantified liver disease following abnormal LFTs.[17–20] Duh *et al.*[17] quantified the incidence of liver enzyme abnormalities in the general population in Massachusetts, USA. There was no long-term follow-up to eventual liver disease diagnosis. A large cohort study in Korea (n = 142,055) reported the association between the LFT serum aminotransferases (aspartate aminotransferase and alanine aminotransferase [ALT]) and mortality from liver disease, indicating that even values that were in the upper quartile of the normal range were associated with worse outcome.[1,18] Our previous study used electronic case record linkage to diagnose liver disease and this project showed that 20% of all LFTs measured were found to be abnormal with less than 10% of these explained by existing liver disease.[21]

To investigate this further we studied the outcomes of individuals up to 15 years post LFT measurement in primary care.[22] Over 60% of the catchment population had LFTs measured. A total of 2,189,152 LFTs were first checked in primary care in 95,992 people. Of these 21.7% were abnormal and 1.26% went on to have a diagnosis of chronic liver disease (hepatocellular causes) while 1.61% developed biliary disease (including gallbladder and biliary tract disease).

Current practice when managing abnormal LFTs in primary care is variable with general practitioner (GP) strategies ranging from ignoring abnormal LFT results, repeat sampling, requesting additional tests or referring to specialist services.[23,24] Recently published guidelines advocate more active investigation of abnormal LFTs, but clinical practice does not yet reflect this.[25]

The intelligent liver function test (iLFT) system was developed to promote appropriate investigation of abnormal LFTs in primary care by utilising minimum diagnostic criteria, the availability of automated tracked analysers and liver fibrosis markers with high negative predictive values.

- A working group convened by the Scottish Government Liver Care Pathway Advisory group used an extensive literature review and expert opinion to achieve professional consensus on minimum diagnostic criteria for liver diseases. Highly specific diagnostic criteria have been identified for each liver disease based on a few simple clinical observations *e.g.* body mass index (BMI), alcohol intake, presence or absence of metabolic syndrome, and blood test results. This allows diagnosis using a minimal range of robust diagnostic criteria but gives confidence that those identified do indeed have the disease. The system is not designed to put everyone in a diagnostic group, just those that meet the criteria. It is designed to fail safe, always defaulting to further clinical evaluation if there is diagnostic uncertainty.[6]
- Technological developments within diagnostic laboratories have led to the use of automated tracked analysers where patient samples are passed between analysers to deliver a wide repertoire of tests under computer control. In real time the system can change a sample's journey based on the preceding results (reflexive testing) *i.e.* if a result outside of a threshold value is detected, the system can use a pre-programmed algorithm to trigger additional tests automatically. Combining results with clinical information provided through the electronic ordering system enables the laboratory information management system to calculate prognostic indices and allocate a suggested diagnosis.
- The crucial point in determining management and need for referral to hepatology for expert review and further investigation for the common liver diseases is the degree of liver fibrosis or cirrhosis, as many with no fibrosis can be treated with lifestyle advice. Non-invasive fibrosis indices are effective at excluding significant fibrosis and many of these indices can be calculated using routinely available clinical laboratory analytes.[26–28]

By integrating these 3 steps we designed a system that enables an intelligent automated response to abnormal LFT results. This study compares this new iLFT system to routine clinical practice in primary care. We hypothesised that iLFT would deliver early identification of treatable liver disease, reduce GP consultations, and be cost-effective.

Patients and methods

Study population

Patients were purposively recruited from 6 general practices in Tayside, Scotland, UK between September 2015 and November 2016 to ensure a mix of urban and rural practices. Each practice has an average of 5 thousand registered patients. Consent was sought from each patient who had LFTs sampled in the intervention group. The inclusion criteria were people aged 18–75 in whom their GP requested LFTs. They were excluded if they had; jaundice, pre-existing liver disease, previously known abnormal LFTs, or LFTs required for monitoring of a specific side effect of a drug or treatment. Review of the post-enrolment clinical records confirmed the absence of exclusion criteria and validated BMI and the presence/absence of metabolic syndrome. Confirmation of alcohol intake was not possible as it was self-reported to the health care professional at the time of consent.

The study was conducted in accordance with the 1975 Declaration of Helsinki and the principles of good clinical practice. The study was co-sponsored by the University of Dundee and NHS Tayside, and was ethically reviewed and approved by the East of Scotland Research Ethics Service.

Study design

A stepped wedge design was used with all 6 participating practices receiving the iLFT intervention. Each practice was randomised to 1 of 3 different start dates, at monthly intervals. Each practice functioned as their own control in a mixed model analysis, over the duration of the trial (6 months control, 6 months intervention) (Fig. S1).

The control population were people with LFTs above the NHS Tayside reference limits (Table S4) in the participating practices during the 6 months before the iLFT intervention. They were retrospectively assessed, all who fulfilled inclusion and exclusion criteria were included as controls. In the intervention arm GPs requesting LFTs during the intervention period could select the iLFT option. All patients with abnormal LFTs were followed up.

Outcomes for all individuals in both groups were accessed via their GP records 6 months after initial LFT sampling, to allow GPs to record and action iLFT recommendations. The final liver diagnosis recorded in the GP notes was extracted and numbers of visits, referrals, and tests performed during the 6 months were recorded. The control period was scheduled first in all practices to avoid confounding by the potential educational effect of iLFT.

The intervention phase worked as follows:

1. GP requests LFTs on their electronic requesting system. A prompt asks if they want to screen for liver disease if LFTs are abnormal.
2. A positive response prompts the GP to enter data about patients' alcohol consumption, BMI and features of metabolic syndrome.
3. If components of the LFT results (bilirubin, ALT, alkaline phosphatase or gamma-glutamyltransferase) are above the NHS Tayside reference limits ("abnormal") this triggers an automated reflexive cascade of additional tests in the laboratory to characterise aetiology. These include viral serology (anti-HCV antibody, HBV surface antigen – positive tests confirmed with PCR), liver immunology (anti-nuclear antibody, anti-mitochondrial antibody, anti-smooth muscle antibody and anti-liver kidney microsomal antibodies), iron studies (ferritin, iron, transferrin and percentage saturation of transferrin), alpha 1 anti-trypsin (with phenotyping if result of <1.0 g/L), and caeruloplasmin. Fibrosis staging algorithms (Fibrosis-4 and NAFLD fibrosis score) are also calculated. These automatically populate the diagnostic algorithms to identify a relevant diagnosis and management plan. There are 31 management plans in total.[6]
4. The report is made available to the GP in real time for them to act on. This included lifestyle advice for ALD/NAFLD patients, management plans for primary care and referral recommendations for those requiring assessment and treatment *e.g.* autoimmune hepatitis and viral hepatitis. Access to the management plans is delivered electronically as web hyperlinks.
5. The study team reviewed patients' notes 6 months post intervention to document the GP recorded diagnosis following receipt of the iLFT outcome and management plan. This "final diagnosis" was adjudicated by the GP with no input from the study team.

Primary outcome

The primary outcome was the rate of diagnosis of liver disease, inclusive of hepatocellular and biliary tract disorders, following detection of abnormal LFTs recorded by the GP.

Secondary outcomes

Secondary outcomes included: the number of GP and patient contacts from the initial LFT sample to diagnosis; the number of referrals to secondary care for diagnosis; cost-effectiveness analysis comparing current clinical practice with the iLFT intervention.

Statistical power

The estimated number of participants required to detect an increase in the diagnosis of liver disease from 1% to 2.5% is 2,658, assuming 80% power and a 5% significance level, for a standard before and after study ignoring the stepped wedge design. The design effect reduces the numbers required because of the multiple steps of repeated measures, despite the effect of clustering.[29] With 6 monthly steps, only 1,284 patients would have to have LFTs to demonstrate a ≥2.5-fold increase in liver disease diagnoses. To ensure adequate recruitment 6 practices were recruited rather than 4, as the loss of a whole practice could jeopardise the trial. An increase in the rate of liver diagnosis to the top end of the published estimates of 10% has little impact but increases the power of the study.

Statistical analysis

This stepped wedge design study includes a period where no clusters were exposed to iLFT. Subsequently at 4-week intervals the clusters were randomised to receive the iLFT intervention. The process continued until all 6 clusters are exposed. Data collection continued throughout the study so each cluster contributes to both control and intervention outcomes.

The analysis of primary and secondary outcomes was carried out using non-linear mixed effect models. The models incorporated fixed terms for intervention (before *vs.* after), time (monthly interval time periods), and accounted for the correlation of patients within practices, and the correlation of repeated measurements over time as random effects using the approach of Hussey and Hughes, 2007.[30] The models were also adjusted for age and gender. Missing data was assumed as missing at random as mixed models have the advantage of allowing for missing data while maintaining the principle of intention to treat, although we did not expect any missing data as the primary outcome is dependent on diagnosis of liver disease which is either present or not. These models also enable adjustment for baseline differences. The primary null hypothesis was that there would be no difference in the rate of diagnosis after introducing the intervention. The analysis incorporated the correlation of patients within practices and of repeated measurements over time. Statistical analysis used SAS9.4.

Economic analysis

A within trial analysis explored the incremental cost per correct diagnosis of the iLFT intervention compared to control (routine clinical practice), at 6 months follow-up from the perspective of the NHS and Personal Social Services. Resource use data (such as GP visits, blood test requests, ultrasounds, fibroscans, and secondary care referrals) were collected during the study for each arm along with the trial primary outcome diagnostic data from the stepped wedge design. Unit cost information[31] was combined with the trial resource use data to estimate the mean cost per patient in each arm for price year 2016. Within study cost-effectiveness was reported as incremental cost per correct diagnosis at 6 months.

A decision analytic model was developed for the lifetime analysis, reporting the discounted incremental cost and quality-adjusted life year (QALY) gains, for cost year 2016, adhering to the National Institute for Health and Clinical Excellence (NICE) reference case.[32] The Markov model extrapolated trial outcomes at diagnosis to account for differences in the impact of ALD and NAFLD on the lifetime costs and QALY of patients receiving the iLFT intervention or routine care.

Recruitment

In the control group, 490 eligible patients with abnormal LFTs were identified and followed up for 6 months after their initial GP appointment. In the intervention group, 229 patients were

Table 1. Demographics, alcohol intake and systolic BP measurement.

Characteristics	Control group (n = 490)	iLFT group (n = 64)
Age, mean [SD]	53 [14.96]	52 [15.05]
Gender % (n)		
Male	55.1% (270)	59.4% (38)
Female	44.9% (220)	40.6% (26)
Body mass index, mean [SD]	30.51 [6.64]	30.46 [3.91]
Alcohol, % (n)		
0 units/wk	32.6% (141)	3.2% (2)
1–21 units/wk	55.7% (241)	84.1% (53)
22–50 units/wk	6.2% (27)	4.8% (3)
>50 units/wk	5.5% (24)	7.9 % (5)
Systolic BP mmHg, mean [SD]	132 [17]	133 [13]

BP, blood pressure; iLFT, intelligent liver function testing.

recruited, 64 (27.9%) had abnormal LFTs. There were no significant differences in age or sex between the 2 cohorts (Table 1). Recruitment continued in the intervention phase for the first 4 practices until the end of the trial in all practices, to increase the number of outcomes. All patients recruited out of the step-wedge time windows were excluded from the primary analysis (Table 2). In the control phase all eligible patients were identified (490), whilst in the intervention phase only those recruited by GPs were included, leading to under recruitment of eligible patients. We calculate that the 64 with abnormal LFTs in the intervention arm were approximately 13% of all likely eligible patients in the practices.

Diagnosis outcomes

Final diagnosis was taken as that recorded in the GP notes 6 months after the test. GPs could ignore or adjust (possibly due to further information from the patient or clinical examination) the iLFT test suggested diagnoses. This was left at the discretion of GPs with no influence from the study team, to replicate real life management of results in general practice. Table 3 shows the concordance between iLFT and GP diagnosis.

Table 4 demonstrates the diagnostic outcomes for the control and intervention arms. The commonest diagnoses were ALD, abnormalities secondary to systemic disease, NAFLD and biliary disease across both groups. "iLFT" diagnoses were additionally refined by the use of fibrosis algorithms to stage disease, *i.e.* ALD unspecified in the control arm becomes ALD with or without fibrosis. Within the fibrosis group 6 patients had advanced fibrosis (Fibrosis-4 >3.25 or NAFLD fibrosis score >0.676).[33] Fig. 1 (and Table S1) shows the major diagnostic category outcomes, defined by final GP diagnosis, unadjusted for step-wedge time windows.

A hepatobiliary diagnosis of any description was documented in the GP record in 56% of intervention cases compared to 16% pre-intervention, demonstrating that the system is effective. The majority of additional diagnoses for the iLFT arm came from the groups labelled "Normalised-no diagnosis" and "Not re-checked-not investigated" in the control arm (Fig. 1). The latter is unsurprising as further investigation uncovers a diagnosis, however for the former, standard practice would not investigate normalised LFTs. Arguably most of these LFTs had not normalised given the controversy over the normal range for ALT.[34]

The study recorded the GP diagnosis as the primary outcome, this did not always agree with the suggested iLFT diagnosis, this discrepancy is shown in Table 3. Overall iLFT supported the GP to a diagnosis in 67% of cases. However, GPs discarded the suggested iLFT diagnosis in 13 cases, providing an alternative diagnosis in 6 cases, or no diagnosis, with no further action taken, in 7 cases. iLFT assigned a diagnosis to 29/64 (45.5%) of the remainder 6 (9%) had fibrosis without aetiology and 29/64 (45.5%) had no fibrosis and no diagnosis. On further note review, the majority of the latter group had features of NAFLD.

The primary outcome analysis of the study was performed only on those patients in the predefined step-wedge time windows. This reduced the control cohort from 490 to 486 and iLFT cohort from 64 to 54. The adjusted difference in rate of liver disease diagnosis is given in Table 5 with a highly significant increase of 43% (95% CI 27–59%, p <0.0002) in the iLFT group compared with controls. For secondary outcomes there were significant increases in rates of visits to the GP pre- and post-diagnosis with relative risk (RR) = 2.00 (95% CI 1.37–2.91) and RR = 3.47 (95% CI 1.63–7.36) respectively. In addition, there was some indication of lower rates of non-liver visits to the GP with RR = 0.77 (95% CI 0.59–0.99), though this only just reached significance. The number of nurse visits and blood requests were not significantly increased. The overall number of visits was not significantly different. Referrals to secondary care were significantly increased with an odds ratio = 8.44 (95% CI 1.99–35.73). It is important to acknowledge that the

Table 2. Study recruitment, flow diagram.

	Control	iLFT
Number of LFTs requested in study sites during trial	12,181	15,150
Number of individuals with abnormal LFTs	490	n.a.
Number of iLFT participants	n.a.	229
Number of iLFT participants with abnormal LFTs	n.a.	64
All cases of abnormal LFTs	490	64
Number of participants included in primary outcome stepped wedge analysis	486*	54*

iLFT, intelligent LFT; LFT, liver function testing.
*Patients excluded as recruited out of time windows.

Table 3. iLFT diagnosis *vs.* GP diagnosis.

	Patients	GP final hepatic diagnosis agreed	GP diagnosis	No diagnosis or no action
iLFT suggested diagnosis	29	16	6	7
iLFT descriptive LFTs with fibrosis	6	–	4	2
iLFT descriptive LFTs no fibrosis	29	–	17	12
Total patients	**64**			

GP, general practitioner; iLFT, intelligent LFT; LFT, liver function testing.
This table shows the number of suggested iLFT diagnoses discarded by GPs and outcomes for patients where iLFT offered a description of LFT abnormality with fibrosis assessment but no diagnosis.

Table 4. Final diagnosis in control and iLFT arm in participants with abnormal LFTs.

Final hepatic diagnosis by GP	Intervention type	
	Control, n (%)	iLFT, n (%)
ALD unspecified	30 (6.1)	0 (0.0)
ALD with fibrosis	0 (0.0)	5 (7.8)
ALD without fibrosis	0 (0.0)	5 (7.8)
Abnormal secondary to biliary disease	15 (3.1)	5 (7.8)
Abnormal secondary to systemic disease	42 (8.6)	7 (10.9)
Acute hepatitis	3 (0.6)	1 (1.6)
DILI	2 (0.4)	2 (3.1)
Gilbert's Syndrome	5 (1.0)	1 (1.6)
HBV	1 (0.2)	1 (1.6)
HCC	1 (0.2)	0 (0.0)
HCV	1 (0.2)	0 (0.0)
NAFLD with fibrosis	1 (0.2)	3 (4.7)
NAFLD without fibrosis	1 (0.2)	9 (14.1)
NAFLD without specification	21 (4.3)	3 (4.7)
Primary biliary cholangitis	0 (0.0)	1 (1.6)
Normalised-no diagnosis	81 (16.5)	1 (1.6)
Not normalised- no diagnosis	72 (14.7)	20 (31.3)
Not re-checked-not investigated	216 (44.1)	0 (0.0)
Total	**490**	**64**

ALD, alcohol-related liver disease; DILI, idiosyncratic drug-induced liver injury; HCC, hepatocellular carcinoma; iLFT, intelligent liver function testing; NAFLD, non-alcoholic liver diseases.

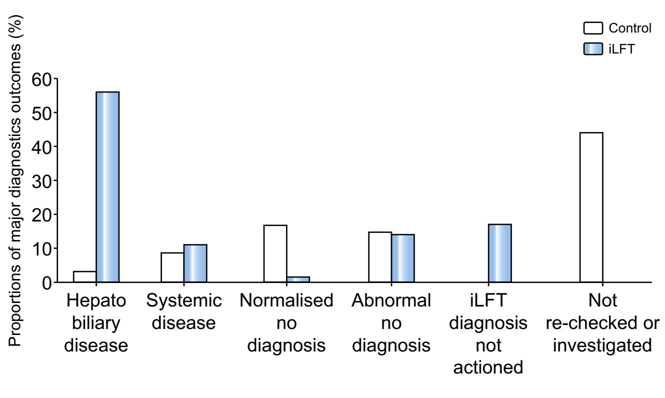


Fig. 1. Proportions of major diagnostic outcomes in all patients with abnormal LFTs in control and iLFT populations. Analysed by linear and non linear mixed effects models. GP, general practitioner; iLFT, intelligent LFT; LFT, liver function testing.

activity reflected in this data in the control cohort is real world practice and not "standard of care" as defined by guidelines, non-investigation of >50% of patients alters the comparator considerably.

Health economic analysis

Table 6 describes the primary trial outcomes from the stepped wedge sample used in the economic analysis, while Table 7 describes the base case economic outcomes. iLFT performs better than control for detecting liver disease (true positive) and identifying healthy people (true negative), resulting in a 51% increase in the probability of correct diagnosis.

The within trial analysis resulted in an incremental cost per correct diagnosis (including true positive and true negative) of £284. In the lifetime model, iLFT resulted in a cost saving of £3,216 per person and improved effectiveness, with an additional 0.021 QALY gained. iLFT is not only cost-effective but a dominant strategy. Fig. S2 presents the distribution of incremental cost and effect outcomes from the probabilistic analysis (1,000 iteration Monte Carlo simulation) on a cost-effectiveness plane. All the simulation outcomes fall in the southeast quadrant. iLFT remains the dominant strategy across a wide range of willingness to pay thresholds, and at the UK threshold of £30,000 per QALY[35] iLFT has a 100% probability of being cost-effective.

GP feedback

GP participants (21/23) completed a questionnaire reviewing iLFT. The majority were positive about iLFT and wished to continue to have access, finding it easy to use, and feeling it reduced their workload (see Tables S2 and S3).

Discussion

The purpose of iLFT was to increase the proportion of patients who underwent appropriate investigation, achieved a diagnosis, and were correctly managed for their condition. The first step of this is clearly to perform the investigations.

Analysis was performed on the final hepatic diagnosis recorded by the GP, rather than the suggested iLFT diagnosis, as it is the clinical decision maker the study was intended to influence. The final diagnosis may well have been suggested by iLFT, however using the diagnosis recorded by the GP as the final diagnosis affords clinician input and reflects the final clinical decision in primary care which determines patient outcome. This biases against the diagnostic rate of iLFT, as several iLFT diagnoses were not actioned or recorded by the GP. GPs have access to additional clinical information that may revise the suggested diagnosis, *e.g.* comorbid disease or changed information *e.g.* alcohol consumption. They may also regard the abnormality as insignificant. The overall GP response to the iLFT diagnostic report demonstrates variance. We think that this was due in part to their taking time to familiarise themselves with the system and its full potential. Another factor is GPs acting appropriately by making a clinical judgement about an individual patient rather than accepting the algorithm generated diagnosis. Familiarity with the system, further education about the reliability, and increased use of iLFT may change some of those behaviours and reduce variation in practice. This system should

Table 5. Primary and secondary analyses of stepped wedge participant outcomes.

Outcomes	Adjusted[†] intervention effect estimate (95% CI)	*p* value
Primary outcome		
Difference in rates (intervention – controls)	43.43 (27.46–59.40)	<0.0002
Secondary outcomes	**RR or OR (95% CI)**	***p* value**
# of GP visits pre-diagnosis (liver)	2.00 (1.37–2.91)	0.0003
# of GP visits post diagnosis (liver)	3.47 (1.63–7.36)	0.0013
# of GP visits (non-liver)	0.77 (0.59–0.99)	0.0496
# of Nurse visits within GP practice[‡]	1.24 (0.72–2.13)	0.4295
# of GP LFT blood requests post baseline[‡]	1.19 (0.71–2.02)	0.5074
# GP visits pre-diagnosis (liver) + GP visits post-diagnosis (liver) + GP visits (non-liver) + Nurse visits within GP practice + GP LFT blood requests post baseline	1.15 (0.98–1.34)	0.0872
# of GP visits pre-diagnosis (liver) + GP visits post diagnosis (liver) + GP visits (non-liver) + Nurse visits within GP practice	1.13 (0.96–1.34)	0.1392
# of GP visits pre-diagnosis (liver) + GP visits post-diagnosis (liver) + Nurse visits within GP practice	1.48 (1.07–2.06)	0.0189
Gastroenterologist referral appts + Endocrinologist diabetologist appts + haematology appts[‡]	8.44 (1.99–35.73)	0.0040

GP, general practitioner; LFT, liver function testing; OR, odds ratio; RR, relative risk.
Analyses based on results 6 months before and 6 months after the intervention.
[†] Adjusted for time, age, gender and alcohol dependence (Linear mixed model for primary outcome and Poisson mixed model for counts).
[‡] Modelled using Negative Binomial.

be utilised as an investigative tool to suggest diagnoses but does not supersede clinical judgement. It is important to note that the clinical activity in the control cohort is real world practice not "gold standard of care" as defined by guidelines. Real world practice is suboptimal resulting in non-investigation of over half of patients.[21]

Importantly, iLFT determines diagnoses in a primary care context allowing for informed and appropriate referral. The iLFT diagnosis was not based on liver histology, nor advanced complex imaging as these modalities do not reflect the early stages of the presentation of abnormal LFTs. Instead the iLFT diagnosis was based on the algorithm that we have validated and published previously using a minimum data set.[6] The consensus process that developed the iLFT algorithm accepted that certain diagnoses would require further investigation, including possible liver biopsy, which would require referral from primary care; the management plans associated with diagnoses were created to provide a fail safe route to expert review.

The most striking effect of the iLFT intervention is the increase in liver disease diagnosis rate. This is attributable to the fact that all LFT's were investigated automatically in contrast to the 50% or more that were not investigated in the control group. In the control group 59% were not actioned further compared with zero in the intervention group. This significantly increased the diagnosis rate by 43% in the intervention group. The benefits of earlier identification of liver disease are evident in the lifetime economic model. As detection of liver disease was higher in the intervention arm, we are certain that liver disease that would previously have been missed was detected. The success of iLFT depends on early diagnosis and health interventions to avert the consequences of missed or late diagnosis of liver disease. The economic model reaffirms this rationale, but longer term follow-up is needed to prove its benefit on overall outcomes.

Increasing the number of liver diagnoses inevitably increases the number of referrals to secondary care. The benefit of iLFT is that the liver diagnoses are stratified to management either in primary care or secondary care. Conditions that require secondary care input are referred appropriately and those that can safely be managed in the community are not referred. The nature of referrals to specialist care are therefore more appropriate and this enables suitable allocation of resources.

Several cases were recommended for referral because of indeterminate fibrosis scores. The fibrosis staging tests used in this system; Fibrosis-4 and the NAFLD fibrosis score, are cheap to use and have well validated cut-offs to exclude significant fibrosis. However, they are not highly specific and many patients scoring above the cut-offs will not have significant fibrosis. In the current system they would default to further assessment in secondary care which adds to cost. Additional tests such as ultrasound or magnetic resonance based elastography may improve fibrosis estimation but would be difficult to automate and the cost would likely be prohibitive. The use of a second line biochemical analyte or scoring system, such as the enhanced liver fibrosis (ELF™) test, which has recently been recommended by NICE for use in NAFLD staging,[36] would be easy to add into the system. This has the potential to reduce costs and referrals by improving fibrosis stratification.

Our model of immediate investigation is in agreement with new versions of international guidelines which previously recommended repeat testing first, as a proportion of LFTs normalise.[25,34] Our analysis shows this is flawed on 2 counts, firstly many of those LFTs that apparently normalised still had underlying liver disease, a proportion of which had significant fibrotic liver disease. Secondly, this project demonstrates that immediate screening for aetiology of liver disease especially on the index sample is highly cost-effective, this is supported in modelling work by Tapper *et al.*[37]

The iLFT algorithm reduces future burdens of liver disease by enabling earlier interventions, guided by fibrosis scores which highlight those most at risk of future liver disease. The impact of this is demonstrated in the lifetime economic analysis which models the pathway for detected and undetected ALD and NAFLD. For the analysis we assumed the increased diagnosis rate reflected these diseases being detected early in the iLFT arm and not investigated further in the control arm. This results in a dominant strategy, lifetime cost savings to the NHS, and overall improvement in QALY gains, due to this earlier detection. iLFT is costlier than current clinical practice due to the additional blood tests, scans, and increase in referrals. The increased diagnostic rate results in an ICER which is considered highly cost-effective. The short-term additional costs are outweighed by the long-term savings to the NHS through earlier identification of ALD and NAFLD. These models are based on

Table 6. Primary diagnostic outcomes using stepped wedge analysis population for Control and iLFT arms.

	Control* n = 486 n (proportion)	iLFT* n = 54 n (proportion)	Probability difference (95% CI)**
Proportion of disease diagnosis (true positive)	120 (0.25)	38 (0.7)	0.45 (0.32–0.59)
Proportion of no disease diagnosis (true negative)	81 (0.17)	12 (0.22)	0.06 (-0.06–0.17)
Proportion of not investigated	213 (0.44)	0 (0)	−0.44 (-0.39–0.48)
Proportion of investigated not normalised	72 (0.14)	4 (0.07)	−0.07 (-0.15–0.005)
Proportion correct diagnosis (true positive & true negative)	**201 (0.41)**	**50 (0.93)**	**0.51 (0.43–0.59)**

*Adjusted for age and sex by probit model.
**Based on bootstrap, 1,000 iterations adjusted for age and sex.

Table 7. Economic outcomes: short term and lifetime iLFT *vs.* Control group using stepped wedge analysis population.

Interventions	Within trial outcome		Lifetime model outcomes	
	Within trial mean cost	Probability of correct diagnosis	Lifetime mean cost	QALY gained
Control	£185	0.41	£59,764	8.523
iLFT	£328	0.93	£56,545	8.545
Difference (95% CI)	£146 (£63–£228)	0.51 (0.43–0.59)	−£3,216 (−£7,643–£897)	0.021 (0.009–0.040)
ICER		**£284 (£128–£440)**		**iLFT is dominant**

iLFT, intelligent liver function testing; QALY, quality-adjusted life years.

conservative estimates of impact of interventions, but there was little uncertainty in the cost-effectiveness results, in almost all modelled scenarios it was the dominant cost saving strategy. We have the direct costs of the process, so the cost-effectiveness analysis is very robust. The estimated impacts of diagnosis of HCV, NAFLD and ALD use established models of disease progression and conservative estimates of intervention impact taken from the literature. Whilst not a replacement for long-term follow-up, this methodology is the most robust available.

The project did not address the reproducibility of this system to other healthcare services, where costs and access to the appropriate technology will vary. However, the systems (or similar systems) used in this study are available globally, so iLFT is likely to be applicable and cost-effective in most systems.

Contrary to expectations there was no significant reduction in GP workload (GP and nurse visits) in the iLFT arm. This can be explained by findings from the control arm, over 50% of LFTs checked were not actioned or followed up. The remaining 50% that were actioned required increased GP workload compared to the iLFT arm, the 50% not actioned or followed up clearly generated no further input. While this pattern of only investigating 50% of abnormal LFTs is clearly cheaper in the short term it is costlier in the long term with detrimental impacts for the patient's quality and length of life. Furthermore, it is suboptimal clinical practice to request an investigation and then ignore the result when it is abnormal.

The number of participants enrolled into the active phase of the iLFT trial by GPs was only a proportion of the total number of LFTs being requested. The study exclusion criteria likely account for a large number of these. Research staff were able to retrospectively identify all eligible patients in the control arm whereas, in the iLFT arm, recruitment required active intervention by the GP. Unfortunately, this does add some potential bias into our results as it is possible that GPs preferentially recruited patients they thought would benefit from the additional testing.

As with any new system, uptake is often slow due to lack of familiarity and time pressures. Additionally, this was overtly a research project, with unknown benefits. In a busy clinical practice, it is often perceived to be easier to avoid involvement in the research. Uptake of the system was higher later in the study. Which reflected the benefit felt by the GP participants, who responded positively to the iLFT system.

A total of 21/23 completed an iLFT questionnaire noting that they wished to continue to have access to iLFT. They felt that it was easy to use and reduced their workload (see Tables S2 and S3). Feedback on the system was very positive and the vast majority of GPs were keen to continue to have access at the end of the study.

Conclusions

Liver disease is increasing in incidence in contrast to many other conditions, predominantly driven by NAFLD. It disproportionately affects people under 65 leading to substantially increased morbidity and mortality. It is clear that interventions that lead to early diagnosis and the opportunity to intervene and abate disease progression are needed. iLFT delivers this opportunity in primary care to the general population at a minimal intervention cost, using existing infrastructure and utilising existing clinical pathways. It is designed for immediate implementation and could have impacts in the short term. The iLFT system works, it increases liver diagnosis, is cost-effective, and is clearly more effective at diagnosing liver disease than the standard of care.

Financial support

This study was funded by the Chief Scientist Office of the Scottish Government. They had no role in study design, data collection, data analysis, data interpretation or writing of the report.

Conflict of interest

The authors declare no conflicts of interest that pertain to this work.

Please refer to the accompanying ICMJE disclosure forms for further details.

Authors' contributions

JFD obtained the funding. JFD, CW, MHM KAB, BB and ED designed the study. JFD, CW, MHM, EMR, PGM, BB and ED contributed to the implementation of the study. AH, PTD contributed to the data analysis and interpretation. KAB and MR did the health economics evaluation. All authors contributed to the writing and approval of the manuscript. JFD had full access to all the data in the study and takes responsibility for the integrity of data and accuracy of the data analysis.

Acknowledgments

Dr Ron Neville Shirley Cleary, Linda Johnston, James Flood, Dr Sarah Inglis, GPs and patients, and Tayside Clinical Trials Unit and Dr David McLernon

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhep.2019.05.033.

References

[1] Donnan PT, McLernon D, Dillon JF, Ryder S, Roderick P, Sullivan FM, et al. Development of a decision support tool for primary care management of patients with abnormal liver function tests without clinically apparent liver disease: a record-linkage population cohort study and decision analysis (ALFIE) Available from:. Health Technol Assess [Internet] 2009;13.
[2] Armstrong MJ, Houlihan DD, Bentham L, Shaw JC, Cramb R, Olliff S, et al. Presence and severity of non-alcoholic fatty liver disease in a large prospective primary care cohort. J Hepatol 2012;56:234–240.
[3] Homer K, Robson J, Solaiman S, Davis A, Khan SZ, McCoy D, et al. Reducing liver function tests for statin monitoring: an observational comparison of two clinical commissioning groups. Br J Gen Pract 2017;67:e194–e200.
[4] Gnanavel S, Hussain S. Audit of physical health monitoring in children and adolescents receiving antipsychotics in neurodevelopmental clinics in Northumberland. World J Psychiatry 2018;16(8):27–32.
[5] Jan Biegus, Hillege Hans L, Douwe Postmus, Valente Mattia AE, Bloomfield Daniel M, Cleland John GF, et al. Abnormal liver function tests in acute heart failure: relationship with clinical characteristics and outcome in the PROTECT study. Eur J Heart Fail 2016;18:830–839.
[6] Miller MH, Fraser A, Leggett G, Macgilchrist A, Gibson G, Orr J, et al. Development and validation of diagnostic triage criteria for liver disease from a minimum data set enabling the 'intelligent LFT' pathway for the automated assessment of deranged liver enzymes. Frontline Gastroenterol 2018;9.
[7] Radcke S, Dillon JF, Murray AL. A systematic review of the prevalence of mildly abnormal liver function tests and associated health outcomes Available from:. Eur J Gastroenterol Hepatol [Internet] 2015;2.
[8] Pratt DS, Kaplan MM. Evaluation of abnormal liver-enzyme results in asymptomatic patients. (Review Article). N Engl J Med 2000;342:1266.
[9] Williams R, Aspinall R, Bellis M, Camps-Walsh G, Cramp M, Dhawan A, et al. Addressing liver disease in the UK: a blueprint for attaining excellence in health care and reducing premature mortality from lifestyle issues of excess consumption of alcohol, obesity, and viral hepatitis. Lancet 2014;384:1953–1997.
[10] Stefano Bellentani. The epidemiology of non-alcoholic fatty liver disease. Liver Int. 2017;37:81–84.
[11] Masarone Mario, Rosato Valerio, Dallio Marcello, Abenavoli Ludovico, Federico Alessandro, Loguercio Carmela, et al. Epidemiology and natural history of alcoholic liver disease. Rev Recent Clin Trials 2016;11:167–174.
[12] Sebastiani G, Gkouvatsos K, Pantopoulos K. Chronic hepatitis C and liver fibrosis. World J Gastroenterol 2014;20:11033–11053.
[13] Pimpin L, Cortez-Pinto H, Negro F, Corbould E, Lazarus JV, Webber L, et al. Burden of liver disease in Europe: epidemiology and analysis of risk factors to identify prevention policies. J Hepatol 2018;69:718–735.
[14] Bouttell J, Lewsey J, Geue C, Antony G, Briggs A, McCartney G, et al. The SCottish Alcoholic Liver disease Evaluation: A Population-Level Matched Cohort Study of Hospital-Based Costs, 1991–2011. PLoS ONE 2016;11 e0162980.
[15] Shearman DJC, Finlayson NDC, Camilleri M, Carter D. Diseases of the gastrointestinal tract and liver. 3rd ed. Churchill Livingstone; 1997. p. 735–752.
[16] Dufour DR, Lott JA, Nolte FS, Gretch DR, Koff RS, Seeff LB. Diagnosis and monitoring of hepatic injury. II. Recommendations for use of laboratory tests in screening, diagnosis, and monitoring. Clin. Chem. 2000;46:2050–2068.
[17] Duh Mei-Sheng, Walker Alexander M, Kronlund Kenneth H. Descriptive epidemiology of acute liver enzyme abnormalities in the general population of central Massachusetts. Pharmacoepidemiol Drug Saf. 1999;8:275–283.
[18] Kim HC, Nam CM, Jee SH, Han KH, Oh DK, Suh I. Normal serum aminotransferase concentration and risk of mortality from liver diseases: prospective cohort study. BMJ 2004;328, 983–983.
[19] Arnold DT, Bentham LM, Jacob RP, Lilford RJ, Girling AJ. Should patients with abnormal liver function tests in primary care be tested for chronic viral hepatitis: cost minimisation analysis based on a comprehensively tested cohort. BMC Fam Pract 2011;12, 9–9.
[20] Stefano Bellentani, Claudio Tiribelli, Gioconda Saccoccio, Marino Sodde, Nicoletta Fratti, Christina De Martin, et al. Prevalence of chronic liver disease in the general population of northern Italy: the Dionysos study. Hepatology 1994;20:1442–1449.
[21] Steinke DT, Weston TL, Morris AD, MacDonald TM, Dillon JF. The epidemiology of liver disease in Tayside database: a population-based record-linkage study. J Biomed Inform 2002;35:186–193.
[22] Donnan PT, McLernon D, Steinke D, Ryder S, Roderick P, Sullivan FM, et al. Development of a decision support tool to facilitate primary care management of patients with abnormal liver function tests without clinically apparent liver disease [HTA03/38/02]. Abnormal Liver Function Investigations Evaluation (ALFIE). BMC Health Serv Res 2007;7, 54–54.
[23] Sherwood P, Lyburn I, Brown S, Ryder S. How are abnormal results for liver function tests dealt with in primary care? Audit of yield and impact. BMJ 2001;322:276–278.
[24] Jansky M, Mattlinger C, Nguyen-Tat M, Galle PR, Lammert F, Jäger J, et al. Abklärung von Leberwerterhöhungen in der hausärztlichen Praxis – Versorgungsrealität in Deutschland. Dtsch Med Wochenschr (EFirst), 2018.
[25] Newsome PN, Cramb R, Davison SM, Dillon JF, Foulerton M, Godfrey EM, et al. Guidelines on the management of abnormal liver blood tests. Gut 2018;67:6–19.
[26] Parkes J, Roderick P, Harris S, Day C, Mutimer D, Collier J, et al. Enhanced liver fibrosis test can predict clinical outcomes in patients with chronic liver disease. Gut 2010;59:1245.
[27] Sumida Y, Yoneda M, Hyogo H, Itoh Y, Ono M, Fujii H, et al. Validation of the FIB4 index in a Japanese nonalcoholic fatty liver disease population. BMC Gastroenterol 2012;12, 2–2.
[28] European Association for the Study of the Liver. EASL-ALEH Clinical Practice Guidelines: Non-invasive tests for evaluation of liver disease severity and prognosis [Internet]. 2015 [cited 2018 Jun 26]. Available from: http://www.easl.eu/medias/cpg/Non-invasive/English-report.pdf.
[29] National Institute for Health and Clinical Excellence. Weight management: lifestyle services for overweight or obese adults [Internet]. 2014 [cited 2018 Jun 21]. Available from: https://www.nice.org.uk/guidance/ph53/chapter/4-Considerations#cost-effectiveness.
[30] Hussey MA, Hughes JP. Design and analysis of stepped wedge cluster randomized trials. Contemp Clin Trials 2007;28:182–191.
[31] Curtis L, Burns A. Unit Costs of Health & Social Care 2016 [Internet]. Personal Social Services Research Unit, University of Kent; 2016 [cited 2018 Jun 21]. Available from: https://www.pssru.ac.uk/pub/uc/uc2016/full.pdf.
[32] National Institute for Health and Clinical Excellence. Guide to the methods of technology appraisal 2013 [Internet]. 2013 [cited 2018 Jun 21]. Available from: https://www.nice.org.uk/process/pmg9/chapter/Foreword.
[33] Dyson JK, Anstee QM, McPherson S. Non-alcoholic fatty liver disease: a practical approach to diagnosis and staging. Frontline Gastroenterol 2014;5:211.
[34] Kwo PY, Cohen SM, Lim JK. ACG Clinical Guideline: Evaluation of Abnormal Liver Chemistries. Am J Gastroenterol 2016 Dec;20(112):18.
[35] Appleby J, Devlin N, Parkin D. NICE's cost effectiveness threshold. BMJ 2007;335:358–359.
[36] Xie Q, Zhou X, Huang P, Wei J, Wang W, Zheng S. The Performance of Enhanced Liver Fibrosis (ELF) test for the staging of liver fibrosis: a meta-analysis. PLoS ONE 2014;9 e92772.
[37] Tapper EB, Saini SD, Sengupta N. Extensive testing or focused testing of patients with elevated liver enzymes. J Hepatol 2017;66:313–319.