UK National Screening Committee

Screening for Hereditary Haemochromatosis

External review against programme appraisal criteria for the UK National Screening Committee (UK NSC)

Ottawa Hospital Research Institute

December 2015

Decisions as to whether criteria were satisfied (met/not met/uncertain) were made solely based on the evidence of the rapid review.

The UK NSC advises Ministers and the NHS in all four UK countries about all aspects of screening policy. Its policies are reviewed on a 3 yearly cycle. Current policies can be found in the policy database at http://www.screening.nhs.uk/policies and the policy review process is described in detail at http://www.screening.nhs.uk/policyreview

Abbreviations List

AMSTAR	Assessing the Methodological Quality of Systematic Reviews
HH	Hereditary Haemochromatosis
KSG	Knowledge Synthesis Group
MA	Meta-Analysis
MRI	Magnetic Resonance Imaging
NPV	Negative Predictive Value
NOS	Newcastle-Ottawa Scale
OHRI	Ottawa Hospital Research Institute
PPV	Positive Predictive Value
SF	Serum Ferritin
Sn	Sensitivity
Sp	Specificity
SR	Systematic Review
TN	True Negative
TP	True Positive
TS	Transferrin Saturation
UIBC	Unsaturated Iron Binding Capacity
UK NSC	United Kingdom National Screening Committee
USPSTF	U.S. Preventive Services Task Force

Plain English Summary

This document reviews the evidence published since 2009 relating to screening for HFE-related Hereditary Haemochromatosis (HH) in adults with no symptoms or family history of the condition. Non HFE-related HH (such as juvenile disease or mutations in the TFR2 and FPN genes) were outside of the scope of this document.

HH is a condition in which iron levels in the body build up over time. Iron is a mineral which is used by the body's cells and is essential for maintaining physical and mental health. Excessive iron levels can cause a number of symptoms such as tiredness (fatigue) and joint pain and stiffness. It can also lead to more serious damage to organ such as liver, pancreas and heart. In some cases it can cause liver cancer.

HH is caused by a fault in a gene called 'HFE'. The fault is inherited from parents and can interfere with the body's ability to control its iron levels. HH is complicated because many people who inherit a faulty gene do not experience any symptoms. This is called 'incomplete penetrance'. This has been considered an obstacle to the introduction of a screening programme because it may lead to people being told they have a problem when this may never happen.

In 2009, the UK National Screening undertook a review to determine whether a programme that screens for HH should be introduced in the UK. In the 2009 report, it was unclear (1) how often and to what degree people with the faulty HFE gene (also known as 'genotypes') develop signs and symptoms and (2) whether there is an effective screening strategy for detecting the disease. These were key issues leading to a recommendation that screening should not be introduced.

This rapid evidence summary searches and evaluates studies for both issues to see if evidence has been produced which would justify a change of policy. The review found a number of problems which would prevent the UK NSC recommending a screening programme.

These were:

- a very small number of studies which tracked the health of people with faulty HFE genes over time to see how many developed symptoms, and the very low quality of included studies means that it was not possible to make conclusions about how often and to what degree people with faulty HFE genes develop signs and symptoms.
- only one study was found that looked at screening for HH. This did not provide an adequate amount of information to make conclusions on whether an effective screening strategy for HH exists.

Because of a lack of good evidence in the review, the UK NSC concluded that a change to the current policy should not be considered at this time.

Executive Summary

Hereditary Haemochromatosis (HH) is a disease which most commonly affects patients of Caucasian ethnicity. Mutations of the HFE genes (C282Y/C282Y, H63D/H63D or C282Y/H63D) along with elevated iron parameters (for example, serum ferritin, and transferrin saturation) characterise the disease. Elevated iron is absorbed by the tissues, potentially leading to damage. HH is a complex disease such that those with the HFE genotype may never develop signs and symptoms of the disease. As such there is incomplete penetrance.

The UK NSC examined screening for haemochromatosis in 2009 and evaluated 15 criteria. As part of its three yearly evidence review process, the UK NSC considered a limited number of key issues identified in previous evaluations.

This rapid evidence summary examined literature published since 2009 which addresses two key questions.

These are:

- What is the penetrance and expressivity of adults who screen positive for C282Y or H63D homozygosity or C282Y/H63D compound heterozygosity?
- Is there an effective screening strategy to identify HH in the general adult population?

The evidence base in these questions has been considered very limited in previous evaluations of screening. The previous 2009 external report for the UK NSC had recommended that a national screening program for HH should not be established in the UK. The purpose of this review is to determine whether any substantive evidence has been published since then which could inform the discussion on screening. It also provides a judgement on whether the evidence suggests that reconsideration of the UK NSC's current policy is warranted at this point.

Key Messages – Penetrance and Expressivity of HFE genotypes

- One low quality systematic review and 12 cohort studies were found.
- The systematic review included 20 studies (published 1998-2008), of which three were longitudinal cohorts of C282Y homozygotes addressing iron overload, amount of iron removed, cirrhosis, fibrosis, elevated AST, elevated ALT, arthropathy, and serum parameters.
- Of the 12 cohort studies, only one was an inception cohort (patients who present without symptoms at baseline or patients who present with early symptoms). Using a modified Newcastle-Ottawa scale, validity assessments ranged from 1 to 4 points (total n=5 or 6). No studies reported on expressivity.
    - **C282Y/C282Y.** One study addressed penetrance (hepatic fibrosis). Other outcomes were reported by mainly by one study each: hepatic iron overload, death due to iron overload, death due to cirrhosis, various cancer, survival, death due to cardiomyopathy.
    - **C282Y/H63D.** One study addressed penetrance outcomes (elevated serum ferritin, elevated transferrin saturation, fatigue). Other outcomes were reported by the same study: liver disease, hepatomegaly, AST/ALT, abnormal metacarpophalangeal, and arthritis.
    - **H63D/H63D.** No studies addressed penetrance.

- The reported rates of penetrance did not differ appreciably to the 2009 evaluation. The amount of evidence was very limited and the quality of evidence for all outcomes was assessed as very low. An insufficient body of evidence in this review period led to the conclusion that the criterion was not met at this time.

Key Messages – Diagnostic Screening Strategies

- No systematic reviews were located. One cohort-type study assessing phenotypic testing followed by genetic testing and compared with liver biopsy for fibrosis and cirrhosis was located. Because the intent of the study was not to evaluate diagnostic accuracy, we were unable to calculate diagnostic accuracy measures. The amount of the evidence was very limited and the quality of evidence was assessed as very low. An insufficient body of evidence in the review led to the conclusion that the criterion was not met at this time.

Implications

On the basis of the evidence considered in this review:

- It was not possible to draw reliable conclusions regarding the penetrance and expressivity of the HFE genotypes in the general population. The quality issues identified in previous systematic reviews (for example the USPSTF, 2004) persist in papers published since 2009.
- It was not possible to determine an effective screening strategy for population screening in regards to diagnostic properties of phenotypic or genotypic screening. Without further studies in this area it is unlikely that the discussion of screening in the adult population will be informed by reliable information.

On the basis of these conclusions, and without further research, the UK NSC concludes there is insufficient evidence to warrant a reconsideration of the current UK screening policy.

Introduction

Hereditary Haemochromatosis

Haemochromatosis is a disease that primarily affects people of Caucasian ethnicity and is characterised by an abnormally high uptake of iron into the organs and surrounding tissues, which can lead to serious complications, such as liver cancer and cirrhosis (1;2).

A large proportion of cases of iron overload are a result of hereditary predisposition. Blood iron levels are regulated by a hormone called hepcidin which is produced by the liver and regulates uptake of iron into cells. Mutations in the HFE gene are thought to cause reduced production and circulation of hepcidin, resulting in greater cellular iron uptake(2). This is referred to as primary or hereditary hemochromatosis (HFE-HH).The homozygous mutation C282Y/C282Y causes about 90% of HFE-HH(3). Other mutations, such as heterozygous compound C282Y/H63D and homozygous H63D, may cause phenotypic expression (iron overload and/or end-stage organ disease), but account for a smaller proportion of HFE-HH in comparison to homozygous C282Y patients(1;2).

Following the European Association for the Study of the Liver (EASL) and the American Association for the Study of Liver Diseases (AASLD) practice guidelines, an individual will be defined as having HFE HH if: (a) they are C282Y homozygous with elevated blood iron stores with or without signs/symptoms; or (b) they have elevated serum ferritin levels and are either C282Y homozygous or C282Y/H63D compound heterozygous(4;5).

The three main stages of HFE-HH progression as outlined by the American Association of the Study of Liver Diseases (AASL)(5):

"***Stage 1*** refers to those patients with genetic disorder with no increase in iron stores who have 'genetic susceptibility.' ***Stage 2*** refers to those patients with genetic disorder who have phenotypic evidence of iron overload but who are without tissue or organ damage. ***Stage 3*** refers to those individuals who have the genetic disorder with iron overload and have iron deposition to the degree that tissue and organ damage occurs."

HFE-HH is a recessive genetic disorder characterised by an overall moderate penetrance (the likelihood of developing signs and symptoms) such that only a limited proportion of the population who possess homozygous risk alleles will actually develop symptoms and fewer still would be expected to develop serious iron overload resulting in end-stage organ disease(5). A U.S. Preventive Services Task Force (USPSTF) systematic review (SR) has highlighted the uncertainty within the literature to provide a precise estimate of penetrance. Based on two included retrospective cohort studies, it was estimated that amongst C282Y homozygotes, 38%-50% developed iron overload and 10%-33% developed HFE-HH related morbidity(6).

Due to the uncertainty about the progression from the genetic mutation to clinical expression of HFE-HH, it has been difficult to standardise the diagnostic criteria for HFE-HH (4). The low and insignificant level of penetrance has also informed discussion of the viability of screening. Based on experience of the HEmochromatosis and IRon Overload Screening study (HEIRS), the authors had suggested not recommending general population screening for detecting HFE-HH, the reason being that the signs and symptoms of HFE-HH does not differ from a control population who does not get screened and does not possess the HFE-HH mutation(3).

Previous UK NSC external review, 2009

The previous UK NSC review, in 2009, identified the USPSTF SR (two cohort studies) and a few additional cohort studies. A few of those studies indicated that less severe forms of the disease (e.g., elevated serum iron) were more frequent than more severe forms of the disease (e.g., liver disease). However, these studies were based on a small number of patients, were not inception cohorts, and dealt only with the C282Y homozygous genotype. A similar pattern was reported in the 13 cross sectional studies in the USPSTF review, and the review concluded that the totality of the evidence base was very limited in terms of study design and the quality of reporting. In general, it is known that the rate of penetrance for the C282Y/H63D genotype is lower(3).

Additionally, the UK NSC review noted that the mechanisms promoting progression from the inheritance of an underlying mutation to expression of clinical symptoms and / or disease were not well understood. The review states that a number of environmental factors appear to modify disease expression, including blood loss from menstruation or donation, alcohol intake, diet and comorbid disease (e.g. viral hepatitis).

The review had also identified two different classes of strategy for HFE-HH population screening (1). The first strategy is **phenotypic**, whereby individuals with elevated blood iron indices are referred to for genetic testing. There was concern that this strategy may lead to under-detection, since those who have the mutation may not show elevated iron indices at the point of screenings(5). The second strategy is **genotypic** where individuals are detected for genetic mutations, and are then referred to for assessment of blood iron indices and other diagnostic processes to determine stage of progression. There was concern that this strategy may lead to over-detection of low-risk patients since the odds of developing clinical signs and symptoms is thought to be very low(2;5).

The review described a small number of papers reporting the use of different biochemical markers, tested at varied cut-off values. Further, patient acceptance reflected a variety of views about the goals of screening in terms of the stage of HH it sought to detect and the outcomes it might see to prevent.

Basis for the current UK NSC recommendation

The 2009 UK NSC external review concluded that “screening for haemochromatosis in the UK general adult population does not currently meet the NSC criteria.”

Several key factors were highlighted by the discussion in the 2009 review:

- a lack of consensus about the goals of screening, for example, whether screening should aim to prevent early symptoms attributable to iron overload, or only serious disease such as cirrhosis.
- a lack of consensus about the overall screening strategy, for example, whether the primary test should aim to detect the genetic mutation or clinical expression of haemochromatosis.
- a lack of randomised controlled trial evidence demonstrating that earlier treatment as a consequence of screening results in better clinical outcomes than later treatment.
- the estimated uncertainty of level of penetrance arising from the mutations which had been reported from cohort studies and the USPSTF SR undertaken in 2005. The review noted the uncertainty in these estimates and suggested that further information on penetrance may emerge from ongoing cohort studies.

On the basis of the 2009 review, the UK NSC recommended that a national screening programme should not be established in the UK.

Current update review

The UK NSC reviews the evidence relating to its policy recommendations every three years. This process is undertaken using rapid reviews for which there are three main purposes:

- to gauge whether there have been significant developments in the evidence base on key questions identified in previous reviews on the same topic
- to establish whether a current recommendation can be reaffirmed
- to establish whether a topic is likely to benefit from further assessment through the development of different types of evidence product, for example: SRs, cost-effectiveness studies, disease modelling exercises or primary research.

Questions for reviews at this stage were derived primarily from the issues identified as problematic in a previous review.

This rapid evidence summary of screening for haemochromatosis in the adult population searched for literature since the 2009 to address two key issues identified in the 2009 review:

- what estimates of penetrance and expressivity exist; and
- what information exists on screening strategies to determine whether an appropriate screening strategy has been identified.

Current rapid evidence summary, 2015

This rapid evidence summary was undertaken by the Knowledge Synthesis Group of the Ottawa Hospital Research Institute (OHRI-KSG), Canada, to identify and summarise the evidence. Rapid evidence summaries are produced using accelerated and/or modified systematic review methods in order to make concessions to accommodate an expedited turnaround time of approximately 12 weeks (7). This rapid evidence summary was guided by an *a priori* protocol developed by OHRI-KSG in consultation with the UK NSC. The protocol allowed for modifications in scope and analysis during the conduct of the rapid review, depending on the amount and nature of evidence that was retrieved. The UK NSC and OHRI-KSG jointly discussed and agreed to these modifications. Decisions as to whether criteria were satisfied (met/not met/uncertain) were made solely based on the evidence of the rapid review.

Key Questions, 2015

The key questions for this rapid evidence summary are:

'What is the penetrance and expressivity of adults who screen positive for C282Y or H63D homozygosity or C282Y/H63D compound heterozygosity?'

and

'Is there an effective screening strategy to identify HH in the general adult population?

a. What are the diagnostic test properties of the genotypic screening strategy for detecting HH?
b. What are the diagnostic test properties of the phenotypic screening strategy for detecting HH?'

Based on the issues identified in previous work the current evidence summary sought to establish whether robust studies are available to inform the discussion. The methods and eligibility criteria are outlined below.

The UK NSC criteria, corresponding key questions, and amount of included literature are summarised in Table 1.

Table 1. Key questions for the rapid evidence summary and corresponding UK NSC policy decision criteria.

UK NSC Criterion (1)	Key Questions	Literature Yield
2. The epidemiology and natural history of the condition, including development from latent to declared disease, should be adequately understood and there should be detectable risk factor, disease maker, latent period or early symptomatic stage.	What is the penetrance and expressivity of adults who screen positive for C282Y or H63D homozygosity or C282Y/H63D compound heterozygosity?	N=13 • Systematic Review (n=1) • Cohort studies (n=12)
6. The distribution of test values in the target population should be known and a suitable cut-off level defined and agreed.	Is there an effective screening strategy to identify HH in the general adult population?	N=1 • Diagnostic accuracy study (cohort-type) (n=1)

Methods

Literature search. Separate searches were conducted for the key questions on 4-5 June 2015. Strategies from the original report(1) were enhanced and tested iteratively by an experienced medical information specialist in consultation with the review team. Strategies were peer-reviewed by another senior information specialist using the PRESS checklist(8). We searched Ovid MEDLINE®, Ovid MEDLINE® In-Process & Other Non-Indexed Citations, Embase on OVID, and the Cochrane Library (Cochrane Database of Systematic Reviews, DARE, CENTRAL, and HTA databases using the Wiley interface (**Appendix 1**) for publication years 2009 to the present. A combination of controlled vocabulary (e.g., MeSH terms) and keywords were used. Vocabulary and syntax were adjusted across databases. Animal-only and opinion-pieces were removed from the results.

Study selection. Search strategy records de-duplicated in Reference Manager were uploaded to the online DistillerSR software (DistillerSR, Evidence Partners, Ottawa, Canada). One reviewer independently screened titles and abstracts; potentially excluded records were verified by a second reviewer. Full text reports of potentially relevant records were screened by two independent reviewers. Disagreements were resolved by consensus or third person review. Screening forms were pilot-tested for each key question before implementation: 10 records for title and abstracts, and 5 records for full text reports.

Data extraction. Information was extracted by one reviewer into DistillerSR. Due to time constraints, 17% of studies were verified by a second person for the penetrance key question. The one study included for the diagnostic screening question was verified by a second person. Disagreements were resolved by consensus or a third person. Pilot testing was conducted on two studies for the 'penetrance' question.

Validity assessment. The 11-point validated AMSTAR tool was used for systematic reviews and interpreted by tertiles of score: 0-3 points as low quality, 4-7 points as moderate quality, and 8-11 points as high quality of conduct(9). Observational studies were assessed using a modified Newcastle-Ottawa (NOS) scale(10). We did not attempt validity assessments on penetrance data of limited utility **(Appendix 2).** A modified Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool was used for diagnostic studies (11) **(Appendix 3)**. One reviewer assessed all reports; 33% of cohort studies for the penetrance question were verified by a second person. Disagreements were resolved by consensus. The one diagnostic screening study was verified by a second person.

Evidence synthesis. For both key questions, the data was too heterogeneous or sparse to attempt a quantitative aggregation of results. Results are described narratively.

GRADE. Domains of the Grading Recommendations Assessment Development and Evaluation (GRADE) framework were used to inform judgements on the quality of the body of evidence for each outcome(12).

Changes from protocol. The NOS scale was modified to suit methodological considerations for the 'penetrance' question. The QUADAS-2 tool, developed for diagnostic test accuracy, was adapted for test concordance analyses. Based on the poor presentation of information in some studies, we included studies which explicitly stated in their methods the outcomes they had sought to address.

Definitions for classifying whether criteria were satisfied
Met: Sufficient amount of quality evidence to confidently estimate an outcome or effect that is unlikely to be changed by further research or conducting a full assessment (SR/MA).
Not met: Insufficient amount of evidence or sufficient amount of evidence of poor quality to confidently estimate an outcome or effect. The estimates of the outcome (a) are likely to be changed by further research, (b) may change if a full assessment (SR/MA) was conducted; or (c) may be substantially different from the true effect. The criterion could also be deemed 'unmet' if the benefits of conducting a SR/MA are unclear.
Uncertain: The constraints of the rapid evidence summary methodology prevent a reliable answer to the question. There is a strong indication that a SR /MA should be pursued.

Eligibility criteria- UK NSC Criterion 2- Key Question

What is the penetrance and expressivity of adults who screen positive for C282Y or H63D homozygosity or C282Y/H63D compound heterozygosity?

Population	General adult population (18 yrs or older) who have screened positive for C282Y homozygosity, H63D homozygosis, or C282Y/H63D compound heterozygosity (HFE-related hereditary haemochromatosis). Non-HFE types of HH (including juvenile disease-type 2a and 2b), and types 3 and 4 (mutations in the TFR 2 and FPN genes) were excluded. Prospective inception cohort studies would provide the most robust evidence, but we did not restrict inclusion to inception cohorts. Inception cohorts were defined as either those who present without symptoms at baseline or those who present with early symptoms of the disease.
Settings	No geographic and setting restrictions
Outcomes (study eligibility not dependent on outcome measures)	• **Penetrance** (% of population with a specific genotype who develop outcome(s)) • **Expressivity** (% of population with a specific genotype who develop range of symptoms/outcomes of a specific phenotype) The following outcomes were relevant: *Biochemical*: reduced/elevated SF, TS, UIBC *Clinical Symptoms:* fatigue, weakness, pigmentation, erectile dysfunction/impotence, and abdominal pain. *End-Stage Organ Disease*: liver fibrosis, cirrhosis • **Other outcomes** *Other Outcomes:* e.g., mortality, cancer
Time-frame	2009 onwards, to identify literature since the 2009 UK NSC report.
Study Design	Systematic reviews (SRs)[A] and cohort studies
Language	English
Publication Type and Status	Full-text articles available to the research team electronically through local institutional subscriptions were included. No grey literature searches were conducted but any unpublished reports were considered for inclusion if retrieved through database searches.

[A] Defined as: (1) at least one database was searched; (2) reported selection criteria; (3) quality assessment of included studies was reported; (4) provided a list of included studies.

Summary of Findings: Penetrance and Expressivity

Literature search results

Of 476 unique retrieved records, 232 of them were considered eligible for full text assessment. Given the paucity of SRs, primary studies were included. One SR and 12 cohort studies met inclusion **(Figure 1)**. It is important to note that some of these studies had also included data on predictors which influence penetrance outcomes; however, those results are beyond the scope of our update and will not be included here (13-15).The records and reasons for exclusion at full-text screening are provided in **Appendix 5**.

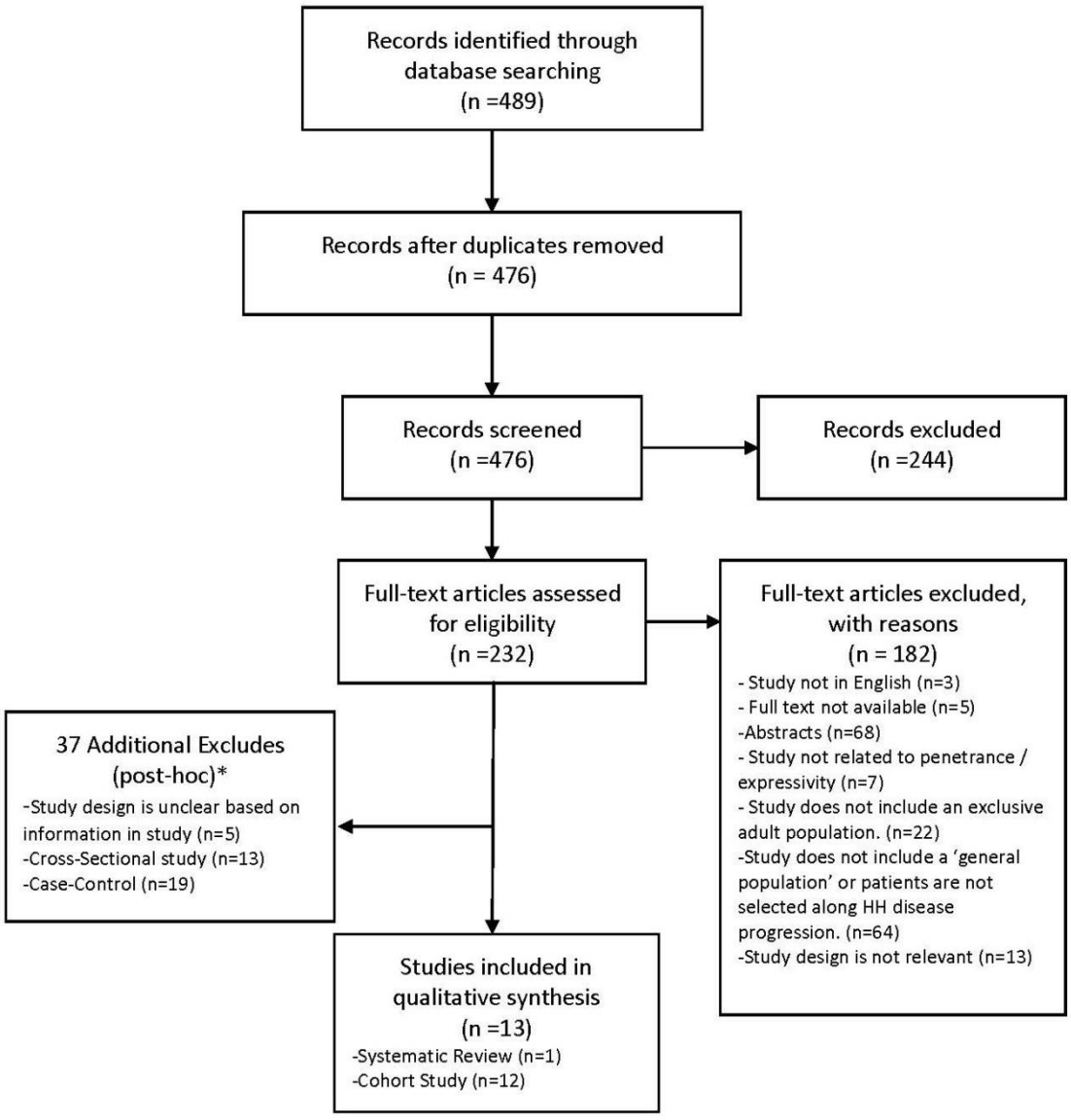

Figure 1. PRISMA flow diagram for penetrance and expressivity.

*See **Appendix 4** for list of 'additional excludes (post-hoc)

Systematic review evidence

One SR conducted by the European Association for the Study of Liver (EASL) was identified but is considered of low quality based on the AMSTAR score of 3 out of eleven (4) **(Appendix 6)**. Authors included 20 studies published between 1998 and 2008. The majority of studies were cross-sectional (n=16; 80%), and remaining studies were either longitudinal cohorts (n=3; 15%), or family-based genetic studies (n=1; 5%). Amongst the three longitudinal cohorts, penetrance of C282Y homozygotes were assessed and included the following outcomes: iron overload (38-50% penetrance), iron removed (42-75% penetrance), hemochromatosis related morbidity (cirrhosis, fibrosis, elevated AST, ALT, arthropathy and serum parameters) (10-33% penetrance).

Primary studies

Characteristics

A detailed table of characteristics for the 12 cohort studies is provided in **Appendix 7.** Three studies were conducted in France and two in Australia. One study was conducted in each of Canada, Italy, The Netherlands, Norway, Sweden, and United Kingdom, and one study was conducted in both Canada and the United States. As a sub-set of the total cohort, the sample sizes of patients genotyped and included in the studies ranged from 18-682. Seven studies provided the duration of follow-up, which ranged from 12 months to 13 years (15-21). Only four studies reported information on ethnicity, which were mostly or completely of Caucasian descent (13;19;22;23). Sex was not reported in two studies (18;19); remaining studies ranged 40% to 75% males.

Only one study was classified as an inception cohort(21), while others presented at later stages of disease(14;15;17;18;22-24) or were unclear regarding disease status at baseline(13;16;19;20). Of those not considered as inception cohorts, 71% of studies had patients who were receiving treatments (e.g., phlebotomy). One study had reported hepatitis C and hepatitis B as co-morbid conditions present at baseline(20).

No studies examined a particular stage of HH disease progression and followed forward to a more advanced stage. For a majority of the studies, staging at both baseline and follow-up was either not reported or varied within the cohort.

Two studies reported a conflict of interest, with authors receiving grants and speakers' honoraria from pharmaceutical industries (18;23). Sources of funding were mostly from governmental agencies, with a few from academia and non-profit organisations. One study was supported in part by industry(17). Two studies had reported funding information; however its categorisation into government, industry, academia or non-profit was unclear (13;22).

Not all studies reported complete information to calculate the penetrance of a given outcome. By definition, penetrance: "…represents the probability of an individual with an affected genotype to manifest the clinical signs of the disease…"(25). Therefore, to calculate penetrance, studies should report the total sample with a given genotype and how many of those who go on to develop signs/symptoms of the outcome. However, some studies only reported the number of genotype patients who developed signs/symptoms (i.e., the numerator). Not reporting the

total sample with the genotype (i.e., denominator) limits the usability of the data. Six studies (across the three susceptible genotypes) exclusively inadequately reported outcomes in this manner, while a seventh study provided some adequately reported information (inadequately reported data provided in **Appendix 7)**.

As depicted in **Figure 2**, of the 12 eligible studies, only 5 provided useable data: 4 studies addressing C282Y/C282Y genotype and one addressing C282Y/H63D genotype. Aside from the 6 studies with inadequately reported data, one study addressing two genotypes (C282Y/C282Y, C282Y/H63D) did not report results separately by genotype to include in our analysis (20).

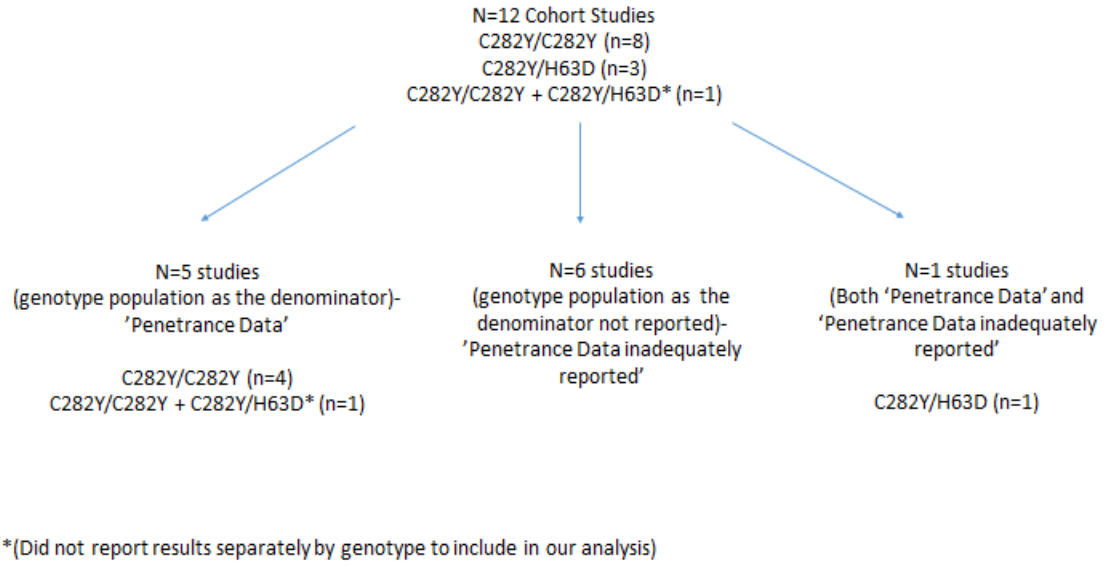

Figure 2. Classification of included studies

Outcomes and Validity assessment

Table 2 summarises the five studies that provide usable information; sample sizes for analysis ranged from 18 to 422, but the largest study was analysed according to two sub-cohorts (American and Canadian samples). Detailed evidence tables for all included studies and associated quality assessment can be found in **Appendix 8** and **Appendix 9**, respectively.

1. C282Y/C282Y GENOTYPE

Penetrance (2 studies). Under our pre-specified outcome categories (biochemical, clinical, and end-stage organ disease), one study of 291 patients who had undergone liver biopsies reported that 44% of patients developed hepatic fibrosis over an unspecified period (**Table 2**)(14). Another study reporting on various biochemical levels only provided median and range data that do not lend to penetrance calculations.

Validity Assessments. Outcomes received two or four points out of five or six for validity assessments using the modified NOS tool (**Table 2- last column**). Objective measures were used to ascertain exposure and outcome status for all penetrance outcomes. For biochemical levels (median and range values reported), studies had adequate duration and little loss-to-follow-up of patients. However, study follow-up and patient attrition were not adequate for the hepatic fibrosis outcome. Studies were not representative of the target population since they were not inception cohorts, and the stage of disease progression was either unclear or mixed at baseline.

Expressivity. No studies reported on this outcome domain.

Other outcomes (1-2 studies). Other outcomes reported on by mainly one study were hepatic iron overload (18.2% of patients), liver cancer (0-1.2%), death due to iron overload (14.8-17.9%), death due to cirrhosis (13.1-17.0%), overall cancer and various subtypes (overall 11.5% to 12.9%; range 1.6-4.7% by subtype), survival (5 year- 67%; mean 12.5 to 13.2 y) and death due to cardiomyopathy (0.9-1.6%).

Validity Assessments. Outcomes received between three or four out of five points for validity assessments using the modified NOS tool **(Table 2- last column)**. Objective methods were used to ascertain outcome status and the completeness of follow-up was adequate. However, only the survival outcome had reported an objective means to genotype patients. As above, studies were not representative of our target population since they were not inception cohorts, and the stage of disease progression was either unclear or mixed at baseline.

2. H63D/H63D GENOTYPE

No studies reported on penetrance or expressivity outcomes.

3. C282Y/H63D GENOTYPE

Penetrance (1 study). Under the pre-specified outcome categories (biochemical, clinical, and end-stage organ disease), one inception cohort of 180 patients reported on biochemical and clinical outcomes, separately for male and female patients, after a mean follow-up of 12 years

(21)(**Table 2**). Eleven percent of females and 37% of males had elevated serum ferritin (SF) levels (>300 μg/L). Twenty-two percent of females and 8% of males had elevated transferrin saturation (TS) levels >45%. Fifteen percent of females compared with 11% of males sought medical attention for fatigue; results were not statistically significant by sex when stratified. Mean biochemical levels were also reported by the same study.

Validity Assessments. Outcomes received between two to four out of five or six points for validity assessments using the modified NOS tool **(Table 2- last column)**. For both types of penetrance outcomes, an inception cohort was used, and exposure ascertainments were obtained through objective measures. A sufficient length of time had passed for the blood iron indices to occur, and the measurement of outcome was objective. For both types of outcomes; however, the extent of patient loss-to-follow–up was inadequate, and authors do not acknowledge whether the outcome was absent at baseline.

Expressivity. No studies reported on this outcome domain.

Other outcomes (1 study). Other outcomes reported by the inception cohort study include liver disease (4-7%), hepatomegaly (1-8%), AST/ALT (1-8%), abnormal metacarpophalangeal joints (20%), and arthritis (2-7%).

Validity Assessments. Outcomes received between two or three out of five points for validity assessments using the modified NOS tool **(Table 2- last column)**. An inception cohort was used. An objective exposure method was used for outcomes. Only AST/ALT and abnormal metacarpophalangeal joints used objective methods to ascertain outcomes. The study did not acknowledge whether the outcomes were absent at baseline, and the degree of losses of follow-up of the entire cohort was inadequate.

Table 2. Population characteristics and evidence summaries for penetrance data (genotype denominator)					
Author; (1) Country; (2) Ethnicity	**(1) Length of follow-up (2)Stage of disease Progression (Baseline→Duration/End of study)**	**(1) Total N (n=x genotyped and included) (2) % Male**	**Outcome**	**Result**	**Quality Score 6 points for TS and SF outcomes; 5 points for all other. Satisfied Criteria¶**
C282Y/C282Y Homozygous					
Wood et al., 2012 (14) (1) Australia; (2) NR	(1) NR (2) Unclear→Mixed	(1) 291; **(n=291)** (2) 67.4%	Hepatic fibrosis	44% (128/291)	2 (B, D)
Bardou-Jacquet et al., 2014 (18) (1) France; (2) NR	(1) Median: 57 months for survival; 60 months for all other outcomes (2) Unclear→Unclear	(1) 736; **(n=18)** (2) NR	SF	Range: 10.7 µg/L -588.7 µg/L (n=13)	4 (B, D, E, F)
			TS	Median±SD: 28% ± 8.5% vs. 68% ± 30%‡, p=0.021 (n=13)	4 (B, D, E, F)
			Serum Iron	Range: 6.14 µmol/L-22.2 µmol/L (n=13)	3 (B, D, F)
			Hepcidin	Median±SD: 13.2 nmol/L ± 8.1 nmol/L vs. 1.61 nmol/L ± 3.6 nmol/L‡, p=0.006 (n=13)	3 (B, D, F)
			Hepcidin/Ferritin Ratio	Median±SD: 6.93 ± 8.75 vs. 0.35 ± 3.6‡, p=0.015 (n=13)	3 (B, D, F)
			Hepatic Iron overload	18.2% (2/11)	3 (B, D, F)
			Survival	**Survival at 1yr after LT:** 83% (15/18) **Survival at 5yr after LT:** 67% (12/18)	4 (B, C, D, F)
Åsberg et al., 2013 (22) (1) Norway; (2) 'mostly Caucasian'	(1)Followed from screening (15 Dec 2009) until date of cancer, dx, death or emigration (2) Unclear→Unclear	(1) 62,860; **(n=292)** (2) 47%	Cancer (overall, colorectal, lung, breast, prostate, liver)	**Overall:** ♀: 11.5% (14/122); ♂: 12.9% (22/170) **Colorectal:** ♀: 2.45% (3/122); ♂: 3.5% (6/170) **Lung:** ♀: 1.6% (2/122); ♂: 1.8% (3/170)	4 –for each cancer outcome (B, C, D, F)

				Liver: ♀: 0% (0/122); ♂: 1.2% (2/170) **Breast:** 2.45% (3/122) **Prostate**: 4.7% (8/170)	
Barton et al., 2012 (13) (1) Canada, The United States; (2) 100% Caucasian	(1) Start date NR (used age at dx). End of follow up was 1 July 2011. (2) Mixed→Mixed	(1) Alabama cohort=294; Ontario cohort=128 **(n=294, 128)** (2) A:63.9%; O: 68.8%	Cause-specific death: Iron overload	***Alabama (SF>1000 µg/L):*** 17.9% (20/112) vs. 3.3 % (6/182)* ***Ontario (SF>1000 µg/L):*** 14.8% (9/61) vs. 3.0% (2/67) § RR (95%CI): ***Alabama (SF>1000 µg/L):*** 5.4 (2.2-13.1)*, p=0.0002 ***Ontario(SF>1000 µg/L):*** 4.9 (1.1-22.0) §, p=0.0359	3 (C, D, F)
			Cause-specific death: Cirrhosis	***Alabama (SF>1000 µg/L):*** 17.0% (19/112) vs. 3.3 % (6/182)* ***Ontario (SF>1000 µg/L):*** 13.1% (8/61) vs. 3.0% (2/67) §	3 (C, D, F)
			Cause-specific death: Cardiomyopathy	***Alabama (SF>1000 µg/L):*** 0.9% (1/112) vs. 0 % (0/182)* ***Ontario (SF>1000 µg/L:*** 1.6% (1/61) vs. 0% (0/67) §	3 (C, D, F)
			Survival	Mean ± SD: ***Alabama:*** 13.2yrs ± 7.3yrs ***Ontario:*** 12.5yrs ± 8.3yrs	3 (C, D, F)
C282Y/H63D Heterozygous					
Gurrin et al., 2009 (21) (1)Australia; (2) NR INCEPTION COHORT	Mean: 12 years (2) Unclear→Mixed	(1) 1,438 **(n=180)** (2) 46.7%	SF	Mean (95%CI): ♀: 120.4 µg/L (100.6-144.0 µg/L) (n=91 post-menopausal at follow up) ♂: 186.5 µg/L (148.9-233.6 µg/L) (n=78)	4 (A, B, D, E)
			SF (>300 µg/L)	♀: 11% (10/90) ♂: 37% (29/78)	4 (A, B, D, E)
			TS	Mean (95%CI): ♀: 38.9% (36.5-41.3%) (n=91 post-menopausal at follow up) ♂: 40.1% (37.1-43.0%)(n=78)	4 (A, B, D, E)

			TS (>45%)	♀: 22% (20/91); ♂: 8% (6/78)	4 (A, B, D, E)
			Fatigue	♀: 15% (14/93); ♂: 11% (9/83)	2 (A, B)
			Fatigue (SF>300 µg/L)	♀: 13% (12/15) vs. 22%‡‡, p=0.45 ♂: 11% (4/38) vs. 12%‡‡, p=0.87	2 (A, B)
			Liver disease	♀: 4% (4/92) ♂: 7% (6/83)	2 (A, B)
			Liver disease(SF>300 µg/L)	♀: 7% (1/15) vs. 6%‡‡, p=0.88 ♂: 8% (3/38) vs. 6%‡‡, p=0.74	2 (A, B)
			Hepatomegaly	♀: 1% (1/77) ; ♂: 8% (5/62)	2 (A, B)
			Hepatomegaly (SF>300 µg/L)	♀: 0% (0/12) vs. 3%‡‡, p=0.61 ♂: 10% (3/10) vs. 3%‡‡, p=0.31	2 (A, B)
			AST/ALT	♀: 1% (1/91) ; ♂: 8% (6/78)	3 (A, B, D)
			AST/ALT (SF>300 µg/L)	♀: 0% (0/15) vs. 2%‡‡, p=0.60 ♂: 14% (5/36) vs. 0%‡‡, p=0.02	3 (A, B, D)
			Abnormal metacarpophalangeal joints	♀: 20% (16/80); ♂: 20% (13/64)	3 (A, B, D)
			Abnormal metacarpophalangeal joints(SF>300 µg/L)	♀: 15% (2/13) vs. 19%‡‡, p=0.76 ♂: 16% (5/31) vs. 23%‡‡, p=0.48	3 (A, B, D)
			Arthritis	♀: 7% (7/96) ; ♂: 2% (2/84)	2 (A, B)
			Arthritis(SF>300 µg/L)	♀: 7% (1/15) vs. 9%‡‡, p=0.77 ♂: 3% (1/38) vs. 3%‡‡, p=0.98	2 (A, B)

‡ Compared to before liver transplant

*Compared to Alabama probands SF≤1000 µg/L

§ Compared to Ontario probands SF≤1000 µg/L

‡‡Normal SF levels (<300ug/L)

¶ Modified New-Castle Ottawa Scale criteria:

A: Representativeness of the exposed cohort; B-Ascertainment of exposure; C-Demonstration that outcome of interest was not present at start of study; D-Assessment of outcome; E- Was follow-up long enough for outcomes to occur?-TS and SF only; F-Adequacy of follow up of cohort

Discussion

Interpretation of Evidence

One SR of low quality and 12 cohort studies were included in the rapid review. Among the 12 cohort studies, a paucity of evidence exists for penetrance and other outcomes for C282Y/C282Y, H63D/H63D and C282Y/H63D genotypes. No studies reported on expressivity.

Using the GRADE framework, the quality of the body of evidence for all outcomes was assessed as very low. All outcomes were downgraded at least one level: **INDIRECTNESS**. Only one study used an inception cohort design. **RISK OF BIAS.** A handful of studies: (a) did not explicitly state the absence of outcomes at baseline; (b) did not conduct blinded or objective assessment of outcomes; and (c) demonstrated inadequate attrition. **INCONSISTENCY**. Because almost all outcomes were represented by one study each, it was not possible to assess for extent of consistency of results. **IMPRECISION**. Due to the nature of the results, we cannot assess for imprecision in terms of the extent of the width of confidence intervals, but sample sizes were generally small. **PUBLICATION BIAS.** Publication (and location) bias is a potential limitation as grey literature searches were not conducted, and five full text articles were not located due to their unavailability within the rapid summary time-frame. The intent of the review was to identify literature since the 2009 UK NSC report, but this assessment does not incorporate the totality of the evidence prior to that.

Reporting issues of encountered penetrance data included studies that reported on combined genotypes (C282Y/C282Y & C282Y/H63D), without providing results separately, and inadequately reporting data to determine the penetrance of outcomes.

Conclusion: Criterion not met to confidently determine estimates of penetrance/expressivity.

At this time, based on the interpretation of the evidence in this review, none of the included studies provide a robust estimate of penetrance/expressivity of the susceptible genotypes (C282Y homozygosity, H63D homozygosity, C282Y/H63D heterozygosity).

The additional literature found since 2009 does not add appreciably to that reported in the 2009 UK NSC review. The 2009 UK NSC review, which includes findings of the USPSTF systematic review (AMSTAR score five, moderate quality), determined that insufficient evidence existed to make a confident estimate of penetrance of iron overload and associated morbidities for C282Y homozygotes. We did identify one inception cohort regarding C282Y/H63D compound heterozygosity, but there is little evidence from which to draw conclusions. As noted by USPSTF, more long-term cohort studies need to be conducted to build the evidence base for more reliable data on penetrance outcomes.

In this rapid evidence summary, 73 full text citations were non-retrievable (68 of which were abstracts, with a large majority being abstracts from conference proceedings). It is unclear at this time how many unique studies those non-retrievable, full-text citations represent and how many would meet our criteria; a brief scan of title and abstracts leads us to believe that we

unlikely to find many that would be eligible. Therefore, an updated systematic review of studies since that undertaken by the USPSTF may not be worthwhile at this time.

The EASL SR (low quality, AMSTAR score of three) included cohort and cross-sectional studies like the USPSTF SR, but does not provide additional information to the above.

More research is needed to determine the penetrance and expressivity across the susceptible genotypes. Without appropriately designed epidemiological studies it is unlikely that a robust estimate of penetrance will be established. Any potential differences in penetrance and expressivity according to sex is of importance to further investigate since it is possibly due to iron reduction in women from menstruation and during pregnancies (5).

Future Research

A core set of outcomes to evaluate in studies should be considered, taking the lead from initiatives such as COMET and OMERACT(26;27). By doing so, we may be able to increase the number of studies focusing on core outcomes, resulting in a greater ability to pool data to provide precise estimates. This may also increase the value of the research.

Eligibility criteria- UK NSC Criterion 7- Key Question

Is there an effective screening strategy to identify HH in the general adult population?

	What are the diagnostic test properties of **genotypic** screening strategy?	What are the diagnostic test properties of a **phenotypic** screening strategy?
Population	General adult population (18 yrs or older) screening context, which included asymptomatic people or a mixed population of symptomatic and asymptomatic people.	General adult population (18 yrs or older), with blood iron indices measured through general health screening. This included asymptomatic people or a mixed population of symptomatic and asymptomatic people.
	Exclusions: Studies which evaluated cascade testing in families of index cases; studies which included solely symptomatic people, as this does not represent a general population screening context.	
Index Test strategy [A]	Screening to detect genetic mutations (HFE) followed by assessment of blood iron indices and other diagnostic processes to ascertain target condition (and stage of progression, as applicable)	Screening to detect elevated blood iron indices (e.g., SF, TF, UIBC) followed by genetic screening (HFE) in screen positive cases.
Comparators: Reference Standard or Other Test (Dependent on screen strategy and stage of HH progression)	As part of an inclusive, sensitive search, we included studies that compared any index test strategy to any other test or reference standard. Below depicts how testing is generally represented as per stage of HH progression. In addition, we included any studies that assessed add-on tests as part of the index test and any variations of a compound reference.	
	Stage 1 [B]: Confirmation by a different type of genetic test. Stage 2 [B]: Different type of genetic test and/or a phenotypic test (SF, TS, UIBC, any combination or differing thresholds). Stage 3 [C]: MRI and/or liver biopsy (for fibrosis/cirrhosis)	Stage 1 [B]: Not applicable for phenotypic testing since stage 1 is genotype only. Stage 2 [B]: Different type of phenotypic test and/or genetic test. Stage 3 [C]: MRI and/or liver biopsy (for fibrosis/cirrhosis)
Outcomes (study eligibility not dependent on outcome measures)	A hierarchy for interpreting the level of evidence based on the acceptability of the reference test as a gold standard and the stage of HH progression: **Higher level of confidence:** -Stages 1 and 2 [B]: test concordance between the index and reference tests. -Stage 3: diagnostic test accuracy measures **Lower level of confidence:** -Stages 1 and 2 [B]: studies report on diagnostic test accuracy without a measure of test concordance -Stage 3: test concordance instead of diagnostic accuracy	
Time-frame	2009 onwards, as an update from the previous UK NSC report	
Study Design	SRs [D] and diagnostic accuracy studies (cohort-type,) [E]	
Language	English	
Publication Type and Status	Full-text articles available to the research team electronically through local institutional subscriptions were included. No grey literature searches were conducted but any unpublished reports were considered for inclusion if retrieved through database searches.	

[A] Considered for inclusion were studies where the sequence of screening strategies could not be determined.

[B] An 'imperfect' reference standard should assess test concordance instead of diagnostic test accuracy(28). Stage 1 considered as is part of the continuum of HH disease progression and monitoring.
[C] An acceptable reference standard to assess diagnostic test accuracy(28).
[D] Defined as: (1) searched at least one database; (2) reported selection criteria; (3) reported quality assessment of included studies; (4) provided a list of included studies.
[E] Cohort-type: one set of inclusion criteria for recruitment

Summary of Findings: Screening Strategy

Literature search results

Of 655 unique retrieved records, 356 of them were considered eligible for full text assessment. No SRs were located, and one primary study met inclusion **(Figure 3)**. The records and reasons for exclusion at full-text screening are provided in **Appendix 10**.

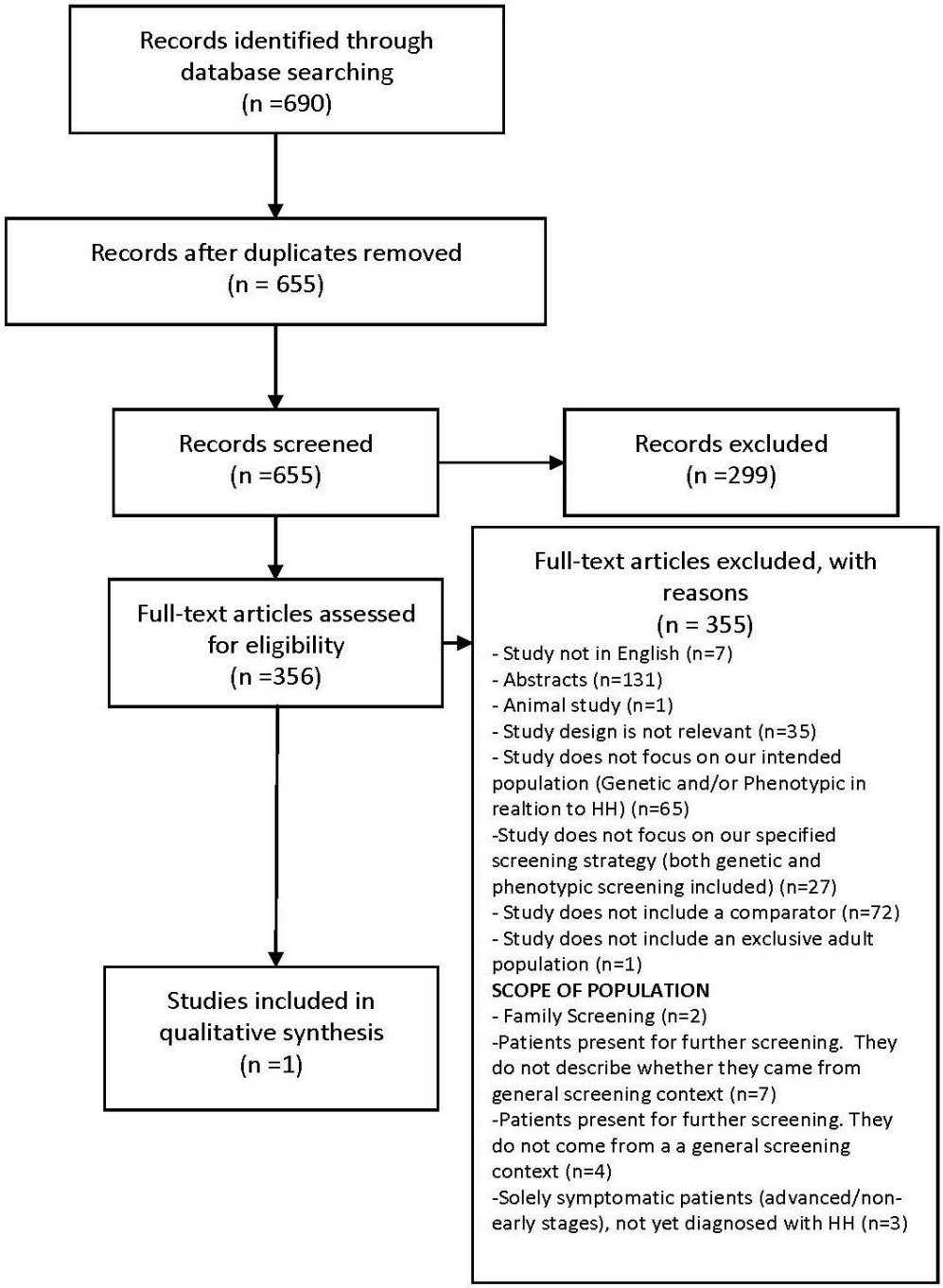

Figure 3. PRISMA flow diagram for diagnostic screening strategies.

Characteristics

The included study was conducted in Sweden and funded by both government and industry(17). The intention of the study was to assess a diagnosed cohort of HH patients and determine their long-time prognosis and survival according to method of detection. We classified this study as a 'diagnostic accuracy cohort-type study', since patients were recruited through a single set of inclusion criteria; however, tests were still conducted cross-sectionally (29). Authors recruited patients diagnosed with HH from nine university hospitals, in which they described the process of index and reference testing to determine the diagnosis. They provided data on a subset of patients identified through routine health check-ups, which satisfied our criteria for inclusion of a general screening population. This subset included a total of 153 patients, of which 79% were male, and the mean age of diagnosis was 45 years. The presence of co-morbidities at baseline was not reported, and none of the subjects were undergoing treatment. Although not explicitly clear it seems that the patients had their blood drawn through a routine check-up, and were then genotyped. Liver biopsies were done at the same time as diagnosis.

Outcomes and Validity assessment

A total of 153 participants with both elevated SF and TS (threshold cut-offs not defined) were included due to detection by their general practitioners or company healthcare providers at routine health check-ups. Without providing exact numbers, the authors had noted that only a portion of the 153 patients had their elevated biochemical levels confirmed by genetic testing, of which those with C282Y/C282Y genotype and an SF level >1000 μg were further given liver biopsies. The remaining subjects had their iron levels confirmed only either by genetic testing or liver biopsy only. Of the 153 patients, only 127 were genotyped (and some further underwent liver biopsy). One Hundred and fourteen subjects were genotyped as C282Y/C282Y; however due to poor reporting, we were unable to determine the genetic make-up of the remaining 13 participants, some of which may consist of genotypes of interest (H63D/H63D, C282Y/H63D) or wild-types.

One hundred and thirty-four (of 153) patients with elevated SF and TS values had undergone liver biopsy (reference standard) at the time of diagnosis (some which were genotyped and some which were not), of which 62 were diagnosed with either fibrosis (n=48) or cirrhosis (n=14).

Diagnostic test accuracy measures (Sn, Sp, TP, TN, PPV, NPV) were not reported by the authors since it was not the intent of the study. Due to missing information necessary to populate a 2x 2 table, we were also unable to manually calculate the diagnostic test accuracy measures.

Quality Assessment

Overall, according to QUADAS-2, the bias associated with this study was considered to be 'high risk'. However, in regards to applicability, the included study matched our pre-specified selection criteria (See **Appendix 11**).

QUADAS 2**								
Domain #1: Patient Selection		**Domain#2: Index test-Genetic**		**Domain#2 Index test-Phenotypic**		**Domain#3: Reference Standard**		**Domain #4: Flow and Timing**
a. Risk of Bias	b. Applicability	a. Risk of Bias	b. Applicability	a. Risk of Bias	b. Applicability	a. Risk of Bias	b. Applicability	a. Risk of Bias
High Risk	Low Risk	Unclear	Low Risk	High Risk	Low Risk	Unclear	Low Risk	High Risk

See **Appendix 11 for detailed quality assessment using QUADAS-2

Discussion

Interpretation of the Evidence

No existing SRs were identified from our search strategy. Only one diagnostic screening study was considered relevant for inclusion.

The one included study addressed a phenotype-then-genotype screening strategy, with liver biopsy as a reference standard for HH. However, only a proportion of patients in the sample had undergone this type of screening strategy. The study did not provide any data regarding clinical utility, but the diagnostic screening strategy was not the main intent of the study. Studies addressing a genotype-then-phenotype screening strategy were not located.

Very little evidence was located regarding genotypic and phenotypic screening strategies, and accordingly, the quality of evidence is very low. **RISK OF BIAS**. The study was either high or unclear for all four domains of QUADAS 2 due to the following reasons: non-inclusion of a random or consecutive sample, not applying the tests to all patients, and poor reporting of pre-specified thresholds for testing. **INDIRECTNESS**. The applicability to our research question for all domains of QUADAS-2 was considered low risk, since the parameters of the diagnostic accuracy study satisfied our inclusion criteria. **INCONSISTENCY**. We were unable to assess consistency due to the inclusion of a single study. **IMPRECISION**. Due to the nature of the results, we cannot assess for imprecision in terms of the extent of width of confidence intervals, but the sample size was small. **PUBLICATION BIAS**. Publication (and location) bias is a potential limitation as grey literature searches were not conducted. The intent of this rapid summary is to identify literature since the 2009 UK NSC report, but this assessment does not include those data.

We did not identify many studies that had addressed diagnostic screening strategies according to our criteria. In terms of liver biopsy studies, it may be difficult to locate general screening population studies since patients who are receiving a biopsy are likely to present with more advanced symptoms and therefore not be considered as part of a general screening context. Further, studies that include liver biopsy may be limited possibly due to its decrease in relevancy of being considered a golden standard(30). Liver biopsy was the main tool used for diagnostic confirmation before the HFE mutation test was widely available. Physicians may be more comfortable performing HFE mutation analysis and phenotypic test as means to diagnose HH as opposed to liver biopsy, since it is less invasive and can identify early stage HH which allows for appropriate treatment(5;30). In terms of stage 1 and 2 patients, no studies have evaluated other types of genetic and phenotypic screening tests in context of the screening strategies.

Reporting issues include incomplete reporting of number of patients undergoing the phenotype-genotype screening strategy, and of those patients, incomplete reporting of how many had undergone reference testing through liver biopsies. As discussed, this was most likely due to the fact that a diagnostic accuracy analysis was not the study's primary intent.

Conclusion: Criterion not met as there is insufficient evidence to determine most effective screening strategy for diagnosing HH in an adult population.

There remains an insufficient volume of suitably designed studies to establish whether there is an effective screening strategy to identify HH in the general population.

Future Research

Irrespective of testing sequence, it remains unclear whether a reference standard exists to confirm the accuracy of stage 1 and stage 2 diagnosis of HH. It also remains unclear whether there are other genetic and phenotypic tests that are available to provide concordance with the current genetic and phenotypic tests for diagnosing stage 1 and stage 2 HH.

Future research should focus on identifying a reference standard to accurately predict the eventual development of hemochromatosis (irrespective of actual development of end-stage organ disease) amongst patients with stage 1 and stage 2 HH, and determining what the subsequent optimal screening strategy sequence should be.

We are unsure as to the merits of conducting a SR. We had identified 131 abstracts (a majority being abstracts from conference proceedings), which were excluded based on our selection criteria for including only full-text articles. It is uncertain as to how many of the full-text papers from the identified abstracts might have been eligible for a SR based on relevancy to the key question. Further, a systematic review would not use study design filters, as was used here. Additionally, relevant studies published before the date of our search strategy (i.e., from the original report) would need to be included in the SR.

Limitations and strengths of the rapid evidence summary

This rapid review was conducted over a condensed period of time (12 weeks). We limited our searching to bibliographic databases and did not search grey literature sources. Although we included only studies written in English, few studies in other languages were located. For the penetrance key question, we were unable to locate five full-text articles. For the diagnostic screening key question, it is generally recommended that study design filters not be used in search strategies developed for diagnostic accuracy SRs; filters were applied in order to manage the literature yield within the timeframe of this rapid evidence summary.

This rapid evidence summary was guided by a protocol developed *a priori*. We first searched for existing systematic reviews before sifting through the primary literature. We had our search strategies peer-reviewed by another senior information specialist using the PRESS form(8). We used standard, systematic approaches for study selection, data extraction, and validity assessment. We also assessed the quality of the body of evidence by adapting the GRADE framework.

Conclusions

Table 9. Conclusions			
UK NSC Criteria being Updated (section x of original report)(1)	**Key Questions being Proposed**	**Conclusions**	
		Overall Conclusion	**UK NSC Criteria was: met, not met, uncertain**
Section 2.The epidemiology and natural history of the condition, including development from latent to declared disease, should be adequately understood and there should be detectable risk factor, disease maker, latent period or early symptomatic stage.	What is the prognosis/penetrance of adults who screen positive for C282Y or H63D homozygosity or compound heterozygosity?	Due to limited evidence, we are unable to make conclusions as to the penetrance and expressivity of HH and other outcomes among patients with the susceptible genotypes (C282Y homozygosity, H63D homozygosity, C282Y/H63D heterozygosity)	Not met -Very low quality of evidence: Insufficient evidence to support key question (one to two studies located per outcome in addition to other issues of quality).
Section 6. The distribution of test values in the target population should be known and a suitable cut-off level defined and agreed.	Is there an effective screening strategy to identify HH in the general adult population? a. What are the diagnostic test properties of the genotypic screening strategy for detecting HH? b. What are the diagnostic test properties of the phenotypic screening strategy for detecting HH?	Due to limited evidence, we are unable to make conclusions regarding the diagnostic test properties of genotype and phenotypic screening strategies to detect HH.	Not met -Very low quality of evidence: Insufficient evidence to support key question (one study located)

Reference List

(1) Allaby M. Screening for Haemochromatosis. A Report to the UK National Screening Committee. Public Health Resource Unit, National Health Service; 2009. Report No.: Revised Report.

(2) Identifying and managing hereditary haemochromatosis in adults. Best Tests 2015;26:14-24.

(3) Adams PC, Barton JC, McLaren GD, Acton RT, Speechley M, McLaren CE, et al. Screening for iron overload: lessons from the HEmochromatosis and IRon Overload Screening (HEIRS) Study. Canadian journal of gastroenterology 2009;23(11):769.

(4) European Association For The Study Of The Liver. EASL clinical practice guidelines for HFE hemochromatosis. Journal of hepatology 2010;53(1):3-22.

(5) Bacon BR, Adams PC, Kowdley KV, Powell LW, Tavill AS. Diagnosis and management of hemochromatosis: 2011 practice guideline by the American Association for the Study of Liver Diseases. Hepatology 2011;54(1):328-43.

(6) Whitlock EP, Garlitz BA, Harris EL, Beil TL, Smith PR. Screening for hereditary hemochromatosis: a systematic review for the US Preventive Services Task Force. Annals of internal medicine 2006;145(3):209-23.

(7) Rapid advice guidelines in the setting of a public health emergency. The WHO Handbook for Guideline Development. 2 ed. 2014. p. 133-55.

(8) Sampson M, McGowan J, Cogo E, Grimshaw J, Moher D, Lefebvre C. An evidence-based practice guideline for the peer review of electronic search strategies. Journal of clinical epidemiology 2009;62(9):944-52.

(9) Grimshaw J, Hill S, Lowe D, Mayhew A, Fiander M, Pratt M, et al. Rx for Change-Methods for Development. CADTH . 2011. https://www.cadth.ca/rx-change.

(10) Wells GA, Shea BJ, O'Connell D, Peterson J, Welch V, Losos M, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. 2007. http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.

(11) Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Annals of internal medicine 2011;155(8):529-36.

(12) Balshem H, Helfand M, Sch++nemann HJ, Oxman AD, Kunz R, Brozek J, et al. GRADE guidelines: 3. Rating the quality of evidence. Journal of clinical epidemiology 2011;64(4):401-6.

(13) Barton JC, Barton JC, Acton RT, So J, Chan S, Adams PC. Increased risk of death from iron overload among 422 treated probands with HFE hemochromatosis and serum levels of ferritin greater than 1000 ug/L at diagnosis. Clinical Gastroenterology and Hepatology 2012;10(4):412-6.

(14) Wood MJ, Powell LW, Dixon JL, Ramm GA. Clinical cofactors and hepatic fibrosis in hereditary hemochromatosis: the role of diabetes mellitus. Hepatology 2012;56(3):904-11.

(15) Manet G, Bardou-Jacquet E, Perrin M, Morcet J, Sinteff JP, Laine F, et al. The iron reabsorption index: a new phenotypic and pathophysiological descriptor in HFE hemochromatosis. European journal of gastroenterology & hepatology 2013;25(11):1321-9.

(16) Adams PC, Barton JC, Guo H, Alter D, Speechley M. Serum ferritin is a biomarker for liver mortality in the Hemochromatosis and Iron Overload Screening Study. Annals of Hepatology: Official Journal of the Mexican Association of Hepatology 2015;14(3).

(17) Aleman S, Endalib S, Stal P, Loof L, Lindgren S, Sandberg-Gertzen H, et al. Health check-ups and family screening allow detection of hereditary hemochromatosis with less advanced liver fibrosis and survival comparable with the general population. Scandinavian journal of gastroenterology 2011;46(9):1118-26.

(18) Bardou–Jacquet E, Philip J, Lorho R, Ropert M, Latournerie M, Houssel–Debry P, et al. Liver transplantation normalizes serum hepcidin level and cures iron metabolism alterations in HFE hemochromatosis. Hepatology 2014;59(3):839-47.

(19) Cobbaert CM, Delanghe J, Boer JMA, Feskens EJM. REGIONAL DIFFERENCES OF HFE (C282Y, H63D) ALLELE FREQUENCIES IN THE NETHERLANDS: A model case illustrating the significance of genographics and prehistorical population migration. Acta Clinica Belgica 2012;67(6):430-5.

(20) Fracanzani AL, Piperno A, Valenti L, Fraquelli M, Coletti S, Maraschi A, et al. Hemochromatosis in Italy in the last 30 years: role of genetic and acquired factors. Hepatology 2010;51(2):501-10.

(21) Gurrin LC, Bertalli NA, Dalton GW, Osborne NJ, Constantine CC, McLaren CE, et al. HFE C282Y/H63D compound heterozygotes are at low risk of hemochromatosis–related morbidity. Hepatology 2009;50(1):94-101.

(22) Asberg ARNE, Thorstensen K, Irgens WO, Romundstad PR, Hveem K. Cancer risk in HFE C282Y homozygotes: results from the HUNT 2 study. Scandinavian journal of gastroenterology 2013;48(2):189-95.

(23) Carpenter JP, Grasso AE, Porter JB, Shah F, Dooley J, Pennell DJ. On myocardial siderosis and left ventricular dysfunction in hemochromatosis. J Cardiovasc Magn Reson 2013;15(1):24.

(24) Aguilar-Martinez P, Bismuth M, Blanc F, Blanc P, Cunat S, Dereure O, et al. The Southern French registry of genetic hemochromatosis: a tool for determining clinical prevalence of the disorder and genotype penetrance. Haematologica 2010;95(4):551-6.

(25) Priori SG, Napolitano C, Schwartz PJ. Low penetrance in the long-QT syndrome clinical impact. Circulation 1999;99(4):529-33.

(26) COMET Initiative. 2015. http://www.comet-initiative.org/

(27) Boers M, Kirwan JR, Wells G, Beaton D, Gossec L, d'Agostino MA, et al. Developing core outcome measurement sets for clinical trials: OMERACT filter 2.0. Journal of clinical epidemiology 2014;67(7):745-53.

(28) Trikalinos TA, Balion CM. Options for summarizing medical test performance in the absence of a "gold standard". Journal of general internal medicine 2012;27(1):67-75.

(29) Deeks JJ, Bossuyt PM, Gatsonis C. Cochrane handbook for systematic reviews of diagnostic test accuracy version 1.0. 0. The Cochrane Collaboration 2009.

(30) Bassett ML, Hickman PE, Dahlstrom JE. The changing role of liver biopsy in diagnosis and management of haemochromatosis. Pathology-Journal of the RCPA 2011;43(5):433-9.

Information Page

Report citation: Quach P, Stevens A, Hersi M, Ghannad M, Ahmadzai N, Garritty C, Moher, D. Screening for Hereditary Haemochromatosis: External review against programme appraisal criteria for the UK National Screening Committee (UK NSC). Ottawa Hospital Research Institute, December 18, 2015.

UK NSC collaborators: John Marshall, Paula Coles

Acknowledgements: The authors thank Becky Skidmore, Roxanne Ward and Raymond Daniel for their assistance with this report compilation. We would also like to thank Dr. Alan Tinmouth for his clinical expertise.

Disclosure of interests of the report authors: The authors declare no competing interests.

Source of funding: The United Kingdom National Screening Committee (UKNSC)