THE UNIVERSITY OF WARWICK

WARWICK MEDICAL SCHOOL

International comparisons of screening policy-making: A systematic review

Farah Seedat, Jennifer Cooper, Louise Cameron, Saverio Stranges, Ngianga-Bakwin Kandala, Hilary Burton, Sian Taylor-Phillips

October 2014

phg foundation
making science work for health

International comparisons of screening policy-making: A systematic review

Farah Seedat,[1]* Jennifer Cooper,[1] Louise Cameron,[2] Saverio Stranges,[1] Ngianga-Bakwin Kandala,[1] Hilary Burton,[2] Sian Taylor-Phillips[1]

*Corresponding Author

1. Department of Health Sciences, Warwick Medical School, University of Warwick, Gibbet Hill, Coventry, CV4 7AL
2. PHG Foundation, 2 Worts Causeway, Cambridge, CB1 8RN

Published October 2014

This report should be referenced as follows:

Seedat F, Cooper J, Cameron L, Stranges S, Kandala NB, Burton H, Taylor-Phillips S. International comparisons of screening policy-making: A systematic review, 2014.

THE UNIVERSITY OF WARWICK
WARWICK MEDICAL SCHOOL
phg foundation
making science work for health

Acknowledgements

We would like to thank all members of our expert panel for their time and information to compile this review: Alison Streetly, Anne Andermann, Anne Mackie, Belen Colomer, Bronwyn Petrie, Catherin Peckham, Frank Schelp, Jackie Spiby, Jane McEntee, Jane Williams, Josephine Taylor, Karen Grimsrud, Kathy Flitcroft, Lars Ehler, Martin McKee, Stacey Carter, Walter Holland.

Contributions

FS created and applied the search strategy to collect the data, and designed and applied the methods to sift, extract and synthesise the data for the search on policy-making processes used by different countries using group B streptococcus as a search term. FS also wrote the report. JC carried out the second review of the article selection, data extraction and data synthesis for the data in this search, and reviewed the paper for redrafting. NBK and SS supervised FS throughout the project and reviewed the report for re-drafting. STP secured the funding, assisted in developing the methods for data collection, synthesising the data and writing the report, and was the lead supervisor for FS. LC and HB used these methods from the group B streptococcus search and re-applied them to genetic screening. LC carried out this search, and sifted, extracted, and synthesised this data. HB supervised LC throughout the project and reviewed the report for re-drafting.

Table of Contents

Acronyms

ASSIA	Applied Social Science Index and Abstracts
Aus	Australia
Can	Canada
CDC	Center for Disease Control and Prevention
Den	Denmark
EC	European Council
EC ELSI	European Commission Ethical, Legal, and Social Implications of genetic testing
ELS	Ethical, Legal, and Social
ESHG	European Society of Human Genetics
Fin	Finland
Fra	France
GBS	Group B Streptococcus
Ger	Germany
GRADE	Grading of Recommendations, Assessment, Development, and Evaluation
HCN	Health Council of Netherlands
HTA	Health Technology Assessment
Ita	Italy
Neth	Netherlands
NNSb	Number Needed to Screen for benefit
NNSh	Number Needed to Screen for harm
NSC	National Screening Committee
NZ	New Zealand
PRISMA	Preferred Reporting Items for Systematic reviews and Meta-Analyses
PHGF	Public Health Genomics Foundation
PPPC	Public and Professional Policy Committee
QALY	Quality Adjusted Life Year
Spa	Spain
SR	Systematic Review
SSCI	Social Science Citation Index
Swe	Sweden
UK	United Kingdom
USA	United States of America
USPSTF	United States Preventive Services Task Force
WHO	World Health Organisation
W&J	Wilson & Jungner

Executive Summary

Background

The UK National Screening Committee (NSC) was founded in 1996 to appraise proposals for new screening programmes, reassess new evidence for screening programmes, and implement and monitor the impact of approved programmes. As part of the NSC's Terms of Reference, a review of their criteria and processes used to assess screening programmes is carried out every three years. In order to assist with this review we conducted two systematic reviews of policy-making processes and criteria used in other countries - one for general and another for genetic screening. We compared these to current UK practice in order to inform this review and future policy-making criteria and processes.

Aims and objectives

The objectives were:

1) To identify evidence relevant to how different countries formulate advice on screening policy, and compare this with current UK NSC practice, including:
    a. Organisations responsible for screening
    b. Topic proposal and selection procedures
    c. Criteria used to assess evidence for screening
    d. Key methodologies used for evidence review and synthesis
    e. Decision-making processes
2) Provide an overview of recent reviews on policy-making used by international bodies or recommended in research literature.
3) To explore any screening criteria and systems developed to assess genetic screening programmes recommended in literature or implemented in practice.

Systems and processes for appraising general screening programmes

(Seedat F, Cooper J, Stranges S, Kandala NB, Taylor-Phillips S)

Methods

To search for policy-making processes for general screening, FS searched Medline, Embase, Applied Social Science Index and Abstracts, and Social Science Citation Index, for articles on group b streptococcus (GBS) screening policies and followed the trail of references to the organisation that made the policy, searching for documentation on policy-making processes for all screening conditions. Therefore the processes found apply to all disease conditions and not GBS alone. We also invited a panel of experts for further identification and searched international health websites. The data was then extracted using a predefined extraction form and synthesised using predetermined tables to aid comparison. Two authors independently conducted the sifting, extraction, and synthesis of the data.

Results

FS and JC examined 670 papers and found recommendations and policy-making processes in 54 papers (figure 1). We found information on policy processes for general screening (and not specific to a condition) for national screening organisations in Australia, Canada, Denmark, Finland, France, Germany, Italy, the Netherlands, New Zealand, Spain, Sweden, the UK, the USA, and for the international bodies, the WHO consultation group on methodology of non-communicable disease screening, and the Council of the European Union. For Belgium, processes concerned GBS only. Policy-making recommendations and reviews were found in 6 research articles (table 1).

Comparing the UK policy-making processes to those of different countries, international bodies, and research recommendations, we found the following areas of difference.

1. Screening systems

All countries, with the exception of Spain, make some degree of national recommendations for screening policy. However decision-making and implementation for screening is not always made at the national level and is delegated to lower level health authorities (table 2). Belgium, France, Germany, the Netherlands, New Zealand, and the UK apply the national screening recommendations to the entire country. Of these countries, national screening organisations only in the Netherlands, New Zealand, and the UK have the national responsibility of implementing the screening programmes, whereas in Belgium, France and Germany implementation is delegated to regional and local authorities. In Australia, Canada, and Sweden, both the decision to screen and its implementation are devolved to lower levels. In Denmark, Finland, and Italy, screening bodies enforce some, but not all screening recommendations, so lower level authorities have to introduce some programmes. Regional and municipal authorities in these countries are able to arrange other programmes not set out in the national screening recommendations. In the USA, the US Preventive Services Task Force (USPSTF) makes national recommendations and in previous years decision-making was left to the regional level and health insurance plans. However, as a result of the Affordable Health Care Act, screening recommendations that are graded A and B by the USPSTF (see appendix) now have to be covered by health insurance plans nationally. This means that health plans must cover USPSTF recommendations to screen, although health plans may still be able to cover screening that is not recommended by USPSTF. In many countries, besides the organisation responsible for screening recommendations, other professional bodies also issue policy recommendations and statements for screening (table 2).

2. Topic selection

In the UK, proposals for new screening programmes from stakeholder organisations will generally be reviewed if an attempt to assess the condition against the screening criteria has been made. In Sweden, Canada, and the USA, topics are suggested by anyone in the public, although in Sweden, an attempt must be made to assess the condition against the first three criteria, which are about the condition. Topics from the list of suggestions are then selected for review in Canada, Sweden, and the USA, by assessing proposals against a list of considerations. In Canada and the USA, a working group is appointed to select a list of topics

to work on during the year. Considerations for deciding topics across the countries include: urgency, need, current practice of condition, important health problem, known natural history, detectable latent stage of disease, timing of most recent review, availability of new evidence, potential impact of recommendations in clinical practice, interest of the public or care providers, variation in care and potential for preventive service to decrease that variation, sufficiency of evidence, new evidence, especially high-quality evidence in a stable field, and potential change to a prior recommendation if there was one.

3. Criteria used for evidence assessment

We found a total of 46 unique items of screening criteria across different countries and research articles (table 3). The criteria not used by the UK, for the most part, are items that are covered broadly but not explicitly mentioned in the UK NSC criteria, and are not left out altogether. The points of difference between the UK NSC criteria and those from a range of other countries, international organisations, and research recommendations can be categorised as follows:

a. Condition – In the UK, a criterion states that the condition should be an important health problem, whilst in other countries the aspects that make a condition important are specified. These were the burden of disease, mortality and morbidity, as well as socio-economic impact.
b. Ethical issues – In line with more than half of the countries included, the UK NSC criteria state that the programme should be ethically acceptable and should have informed consent. Denmark expands on this to specifically include consideration of the consequences of false positive and false negative test results, overdiagnosis, and stigmatisation. Other countries including Australia, France and the Netherlands explicitly consider the promotion of human rights, confidentiality, autonomy, equity and access. Ethical issues often arise around people with indeterminate, intermediate, sub-clinical or carrier status results from screening, for which Canada and the USA have a criterion about knowing the strength of association between intermediate and clinically relevant outcomes, and Australia require a policy for the management of such high risk individuals.
   The UK alone addresses the issue of carrier status with criteria stipulating we should understand the natural history of carriers, and screening should be acceptable to carriers. Both the UK and Spain include criteria that for genetic mutations the subset reported should be clearly stipulated.
c. Measures of Effectiveness – Whilst Australia, Canada, the Netherlands, Spain, Sweden, and the USA, all require evidence of effectiveness in reducing morbidity and mortality, the UK extend this to specifically require high quality randomised controlled trial (RCT) evidence, New Zealand ideally requires RCT evidence, and France requires RCT evidence or international consensus. Canada, Sweden, and the USA may fulfil this need with the use of the Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group for quality appraisal or similar tools (see evidence synthesis methodologies section). Whilst five countries including the UK stipulate that there should be an effective treatment, Spain, specifies that treatment must reduce premature mortality or improve quality of life, whereas Blancquaert et al. (2009) argue

that for genetics the intervention need only offer informed reproductive choice. None of the countries as far as can be identified use Blancquaert et al.'s recommended criterion.

d. Planning and implementation considerations - The planning, implementation, and resource requirements are more explicitly considered in the criteria in Australia, Canada, Spain, and the Netherlands. Criteria in these countries require a recognised need, objectives, and specified target population as well as the inclusion of patient values and preferences in the policy-making process. The NSC criteria currently includes "adequate staffing and facilities for testing, diagnosis, treatment and programme management" and these other countries explicitly add extra considerations of the organisational aspects of programmes such as the optimal screening interval; infrastructure needs; training and education; and the need for databases of population registers and data collection systems.

   The UK alone explicitly requires consideration of the public pressure for widening the inclusion criteria, and the UK and Spain both consider optimising clinical management prior to screening. Cost effectiveness was the most widely adopted criterion across countries, and the UK, Spain, and France, also consider all cost-effective primary preventions prior to commencing screening.

e. Test performance – Different wording is used across countries to describe the desirable test performance. In the UK it is simple, precise, safe and validated, whereas other countries consider a variety of other criteria and wording including; accurate, sensitive, specific, high negative and positive predictive value, reliable, reproducible, efficient, and usable on a large scale.

   Only the UK and Australia ensure that the distribution of test values in the target population should be known and a suitable cut-off level defined and agreed.

f. Quality assurance – Italy specifies that activity of early diagnosis done outside of organised screening programs should be subjected to quality control, and the Council of Europe recommends that if quality assurance standards are not met in the long term, it should be possible for the screening programme to be corrected, and, if this is not possible, stopped. As far as can be identified none of the countries apply the Council of Europe recommendation.

4. Evidence synthesis methodologies

As with most other countries, the UK generally uses the same methods of systematic reviewing. In the UK, the USA and Canada, first the type of review is decided – in the UK the decision is between a knowledge update or full systematic review, and in the USA and Canada it is between a full, target, (on selected questions) or staged systematic review (answering selected questions in an orderly fashion). Compared to the UK, in some countries including the USA and Canada, the procedures within the evidence review are often standardised (table 4). These are:

a. Quality appraisal - The UK leaves the quality appraisal tool to the external reviewer, who utilises the most appropriate tool to assess the quality of individual studies according to the research question. Canada, France, Sweden and the USA utilise prescribed quality appraisal tools. Canada and Sweden use the methods of the GRADE

Working Group where the quality of evidence is assessed for the combination of studies for each important outcome in question and then across the outcomes to score the body of evidence for the whole programme. The USA uses similar methods developed by the USPSTF.

b. Evidence for each criterion – We found tools developed by Harris et al. (2011) and Blancquaert et al. (2009) on the types of evidence required to assess each outcome/criterion. We did not find these tools applied in any of the countries as far as was identified.
c. Calculation of net benefits – A synthesis tool used by the USA and described by Harris et al. (2011) is the magnitude of net benefits. Using the evidence found on the benefits and harms, the magnitude of harms is subtracted from the magnitude of benefits to calculate the net benefit. The USA grades the net benefit into substantial, moderate, small, and zero/negative grades. Whether the net benefits justify the resources is then considered. Harris et al. provide a detailed manual on the approaches and methods to calculate net benefit.

5. Decision-making processes

Similar to the UK, screening committees in the various countries come to a decision about whether to recommend screening or not through discussion in their meetings. In some countries, there are stipulated processes during decision-making. These are:

a. Voting system – Canada, Germany, and the USA apply a voting system to make the final decision on a screening recommendation. In Canada and the USA a quorum is set for the level of agreement required – in Canada this is two-thirds of the members and in the USA it is 10 out of 16 votes.
b. Strength of recommendation – In addition to grading the quality of evidence, Canada, Sweden, and the USA also grade the strength of each recommendation that is made by the screening organisation. For example, using the GRADE system in Canada, recommendations are graded as strong or weak based on factors that include the effect size of the intervention, quality of the evidence, and patients' values and preferences. This process may be more applicable to contexts where the decision to screen is not made centrally. For topic selection purposes, there is the possibility to link the strength of a recommendation to the timing of the next review for a recommendation.
c. Decision support guide - Blancquaert et al. (2009) have developed a decision support guide in which their criteria are separated into three decision-nodes representing the perspective of the individual, family, and society. This is to provide the opportunity to make the perspectives explicit in the discussion and decision of whether benefits outweigh harms. It should help to settle the potential tensions and trade-offs between the perspectives and ensures a proper balance between them.

Systems and criteria for appraising genetic screening programmes

(Cameron L, Burton H)

Methods

LC searched four databases: Medline, Embase, ASSIA and SSCI, for articles on genetic screening programme appraisal using a modified search strategy based on the search terms set out by FS and STP in their review in addition to adding articles from hand-searching. One author reviewed titles and abstracts to identify papers, which were focused on genetic screening programmes. One author extracted and synthesised the data using a customised extraction and synthesis form.

Results

The searches identified 3852 articles, of which 87 papers were potentially eligible for inclusion. Along with papers identified through hand searching, a total of 6 papers were identified which were relevant to genetic screening criteria and systems.

1. Screening systems

Generally, decision-making for genetic screening is covered by the same organisations that develop guidelines for general screening. However in the Netherlands the system for general screening is supplemented by a specialist committee on genetic screening (table 5).

2. Criteria used for evidence assessment

There were criteria items that are covered broadly but not explicitly mentioned in the UK NSC criteria, and are not left out altogether and a further three criteria that were not mentioned in the NSC list (table 3).

a. Criteria not explicitly covered by the NSC

Condition – The Public Health Genomics Foundation (PHGF) (2010) recommend that for genetic screening, particularly for rare diseases, the overall burden of disease should be considered and the importance of the condition may need to be assessed using a global rather than a piecemeal approach assessing each disorder separately. Goel (2001) suggests that the level or risk assessed through penetrance is an important consideration in appraising genetic screening programmes.

Ethical issues - The European Society for Human Genetics (ESHG) Public and Professional Policy Committee (PPPC) recommends that screened patients should be considered before relatives, and comments that economic criteria alone should not be used to justify a screening programme.

Evidence for screening - Bonham (2013) and PHGF (2010) recommend that alternative study designs such as observational cohorts should be considered in place of RCT evidence, which due to the rarity and heterogeneity of genetic conditions, is impossible to provide.

Planning and implementation considerations - Bonham (2013) indicated that these resource criteria can be difficult to satisfy for genetic conditions as it is often the implementation of a screening programme which acts as the impetus to optimise clinical management of cases and provision of facilities.

b. Additional criteria
ESHG PPPC recommends additional criteria not covered by the UK NSC. These criteria include; (1) caution against introducing screening because it is technologically possible, but waiting until there is enough consistency in the characteristics of the conditions screened to allow properly informed consent from the consumer, (2) subjecting cascade screening to the same procedures as population screening, and (3) in populations composed of sub-populations with different genetic backgrounds, selecting the tests according to population substructure and considering focused screening programmes in the sub-populations if the community agrees.

c. Criteria unique to NSC
Three of the five criteria that were unique to the NSC, and not covered in other genetic criteria lists, focused on genetic screening programmes. These were (1) the needs of genetic mutation carriers, in terms of the natural history of this group and in particular the psychological implications of screening, (2) the importance of defining and explaining the subset of mutations which will be tested for if this does not cover all known mutations and (3) the importance of acceptability of the programme for carriers of a mutation and their families.

Conclusions

The UK is viewed quite favourably in the literature for their screening system and policy-making processes in the assessment of screening programmes. We have identified the differences in the criteria and tools used in other countries compared to the UK, which can aid the UK NSC in choosing the areas of focus for their policy review. The systems and processes identified in this review can assist policy-makers in other countries to safely implement health screening programmes within their contexts.

Introduction and Background

The UK National Screening Committee (UK NSC) is argued to be the most significant development in the screening field and a model for organising screening worldwide.[1] It has been praised for having an excellent process for developing evidence-based policy and practice.[2] In deciding whether to start a screening programme, the UK NSC undertakes a policy review, which assesses the proposed screening programme against a set of criteria to ensure the programme's viability, effectiveness, appropriateness, and ultimately to judge whether benefits of a screening programme outweigh the harms. The criteria are based on Wilson and Jungner's recommendations in 1968,[3] that included considerations of the condition, the availability and acceptability of treatment, a policy on who to treat, screening facilities, a suitable and acceptable test, and screening to be continual. The step-wise review process used by the UK NSC begins with identifying relevant stakeholders and moves on to commissioning an independent evidence review to inform the assessment against the screening criteria, an external consultation and finally a meeting to discuss and finalise a decision.[4] Policies are updated on a three yearly basis whereby stakeholders are engaged, knowledge is updated, and if deemed necessary, an external review conducted.[4]

Processes used in other countries to decide whether or not to start or continue a screening programme differ to the UK. Countries have accepted the screening criteria based on Wilson and Jungner's report,[5] but have adapted it and utilise it in different ways.[2] Some differences reflect the different health systems and structures within countries.[1] Other differences reflect developments that have occurred in medicine and healthcare, such as consumer choice, quality assurance, accountability, evidence-based healthcare, and the advent of genetic screening.[6-9]

A comparison of the processes across countries would be useful to develop our understanding of different practices. It would also ensure that policy-makers are operating to the most robust evidence base and processes available internationally. The UK NSC criteria were last updated in 2009 in a response to a Department of Health request to reflect fully the need for thorough economic analysis. The UK NSC is currently reviewing the criteria and processes used to decide whether or not to implement a screening programme. As part of this, we conducted a systematic review of policy-making processes and criteria used in other countries. To meet the objectives of this review, two complimentary methods were applied separately; one to investigate the systems and policy processes used for general screening programmes and another to investigate the systems and criteria used specifically for genetic screening programmes.

Aims and objectives

Aim

Our aim was to explore the health systems, policy-making processes, and criteria used or recommended across the world to formulate screening policy, comparing these to current UK practice.

Objectives

The objectives of the review were:

1) To identify evidence relevant to how different countries formulate advice on screening policy and compare this with current UK NSC practice, including:
    a. Organisations responsible for screening
    b. Topic proposal and selection procedures
    c. Criteria used to assess evidence for screening
    d. Key methodologies used for evidence review and synthesis
    e. Decision-making processes

2) Provide an overview of recent reviews on policy-making used by international bodies or recommended in research literature.

3) To explore any screening criteria and systems developed to assess genetic screening programmes recommended in literature or implemented in practice.

Systems and processes for appraising general screening programmes

Seedat F, Cooper J, Stranges S, Kandala NB, Taylor-Phillips S

Introduction

This chapter describes a systematic review to determine how different countries formulate advice on general screening policy, and how they compare with current UK NSC practice. This includes the organisations responsible for screening, the procedures for topic proposal and selection, the criteria used to assess evidence for screening, the key methodologies used for evidence review and synthesis, and the decision-making processes. The review also examined any recent reviews of policy-making processes used by international bodies or recommended in research literature.

Methods

Search strategy

This search was undertaken using a predetermined protocol that followed the PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) guidelines.[10] In order to find information on the criteria and processes that countries currently use to assess screening programmes, we needed to search for information that is not typically available in scientific journals. Therefore we searched published literature on policies developed for the screening of one condition and then followed the trail of references to the policy-making organisation and the processes used by them. We chose group B streptococcus (GBS) as the condition due to its recent consideration for implementation in several countries.

One author conducted a systematic literature search on 29 November 2013 on electronic bibliographic databases Embase, Medline, Applied Social Sciences Index and Abstracts (ASSIA) and Social Science Citation Index (SSCI) on Web of Science, from 1996 to 2013. The year 1996 was the first that GBS screening was proposed in the USA and later followed by other countries.[11] We developed and applied a predefined search strategy combining the terms for 'group B streptococcus' and 'screening' and 'policy' (see appendix).

Two authors (FS, JC) independently reviewed titles and abstracts and included articles that met the following predefined criteria: 1) population-based screening was the intervention, 2) GBS was the condition 3a) the paper addressed policy-making processes or 3b) stated the organisation that made screening recommendations. We searched references of the abstract-included papers, and websites for the screening organisations, for papers with policy-making information using criteria 1 and 3a. The two authors then independently reviewed the full text of the abstract-included papers, and included those that met the first two criteria and had adequate information about policy-making processes (3a). We excluded articles that were published before 1996, did not contain sufficient information on topic

selection, criteria used, evidence review methodology, or decision-making processes, or were abstracts or letters. We informally translated papers in foreign languages. Any discrepancies between authors' decisions were discussed and were agreed. A third author (STP) arbitrated when consensus could not be reached.

In addition, we invited the participation of experts on screening policy from around the world to identify any useful documentation on policy-making processes in countries they were familiar with or policy-making recommendations in the literature. Panel members are listed in the acknowledgements. Finally, we also searched databases of the European Council (EC), the European Commission, the European Observer, and the World Health Organisation (WHO), for policy-making recommendations.

Data extraction and synthesis

Due to the descriptive nature of this review, it was not necessary to assess the included papers for quality. Two authors (FS, JC) independently extracted data from all included papers on a pre-specified form that included, author, title, year, condition, country, level of recommendation, screening organisation and authority, topic proposal and selection, screening criteria, stakeholder involvement, evidence synthesis methodologies, and decision-making processes. The two authors (FS, JC) then synthesised the data extracted by comparing information for each country or research recommendation, and separating the information into screening organisation, topic proposal and selection, criteria used, evidence review and synthesis methodology, and decision-making processes. We tabulated the details of the policy-making process in separate sections to aid comparability. Based on preliminary work, we anticipated finding other organisations, such as professional medical bodies which make screening recommendations, but which are not screening organisations themselves. We mention the organisations in the results but do not examine their processes.

To systematically examine screening criteria, we listed each criteria item used by the UK NSC[4] and each criteria item developed by Wilson and Jungner.[3] We then analysed the use of these items for each country or research recommendation. As additional criteria were found, we added them onto the list and assessed each for inclusion in the remaining countries or recommendations. Only the information used in the statement of each criterion, and not the details about the criterion, was inputted into the data synthesis table. We also decided that for countries not explicitly using criteria, we would extract information from their processes. Similar to article selection, discrepancies in the authors' extraction and synthesis were discussed, and another author (STP) arbitrated when agreement could not be reached.

Results

Flow of information

The search strategy resulted in a total of 436 unique hits. Scrutinising their titles and abstracts resulted in 171 articles, 5 of which were included in the final analysis.[12-16] The reference check of 171 articles resulted in finding another 31 papers.[4,17-46] We found 1 additional article from international websites,[47] and the expert panel provided a further 17 papers.[1,2,5-9,48-57] A flow chart describing the selection of studies is shown in Figure 1.

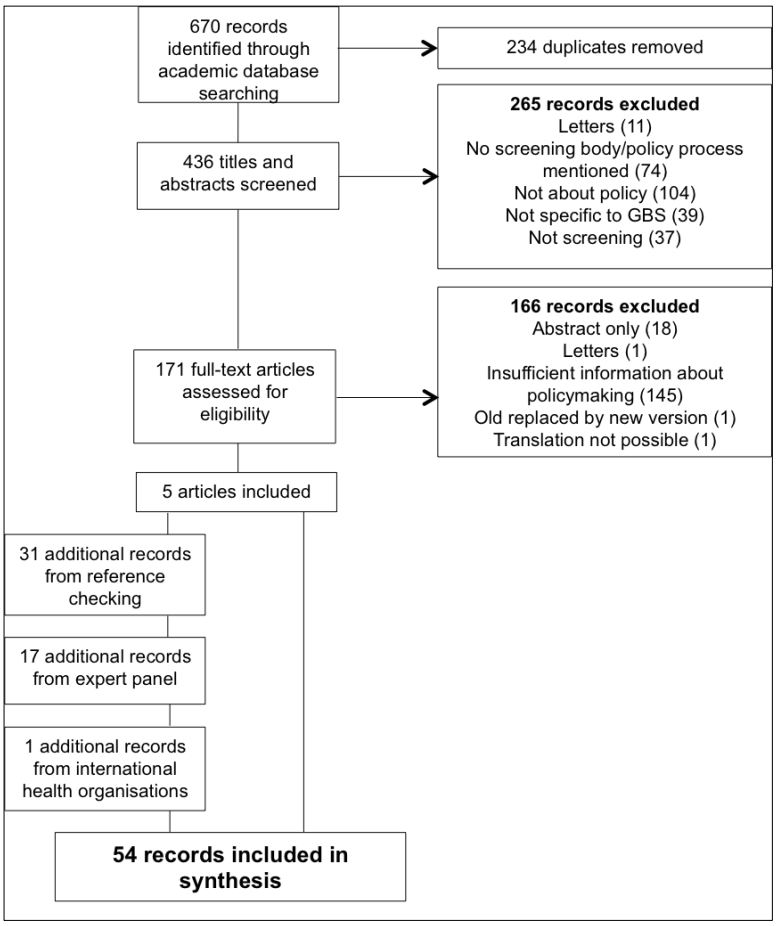

Figure 1: Flowchart of information from the general screening search

General characteristics of information found

From 54 papers, we found information on Australia,[23,24,26,27,52] Belgium,[12,15,22] Canada,[13,21,44,45] the Czech Republic,[17] Denmark,[48,51] Finland,[38] France,[19,35] Germany,[40,41] Italy,[14,36] Japan,[28] the Netherlands,[25,32,33] New Zealand,[18,34,53] Spain,[29,42] Sweden,[55] Switzerland,[31,43] the UK,[4] the USA,[16,20,30,39,46] the WHO Consultation Group Report on methodology of non-communicable disease screening,[37] and the Council of the European Union.[48,50] We also found summaries on policy-making across various countries and institutions[1,2,5,6,49,56,57] as well as research recommendations[7-9,47,54] (see table 1).

Table 1: General characteristics of information found

Name	Conditions found	Screening organisation	Criteria used	Topic selection	Evidence review	Decision-making
COUNTRIES						
Australia[23, 24, 26, 27, 52]	Cancer, Group B Strep, General, Neonatal sepsis	Y	Y	N	R	R
Belgium[12, 15, 22]	Group B Strep	Y	N	Y	R	Y
Canada[13, 21, 44, 45]	General, Group B Strep	Y	NU	Y	Y	Y
Czech Republic[17]	Group B Strep	N	N	N	N	N
Denmark[48, 51]	General	Y	Y	R	Y	N
Finland[38]	General	Y	Y	N	Y	N
France[19, 35]	General, Early neonatal bacterial infection	Y	Y	Y	Y	Y
Germany[40, 41]	General, GBS	Y	Y	Y	R	Y
Italy[14, 36]	Cancer, Group B Strep, General	Y	Y	N	R	N
Japan[28]	Obstetrical practice	N	N	N	N	N
Netherlands[25, 32, 33]	General, Group B Strep	Y	Y	Y	Y	R
New Zealand[18, 34, 53]	General, Genetic, Group B Strep	Y	Y	Y	Y	N
Spain[29, 42]	General, Group B Strep	Y	Y	N	Y	N
Sweden[55]	General	Y	Y	Y	Y	Y
Switzerland[31, 43]	General/ Group B Strep	NU	NU	NU	NU	NU
United Kingdom[4]	General, Group B Strep	Y	Y	Y	Y	Y
United States of America[16, 20, 30, 39]	General, Group B Strep	Y	NU	Y	Y	Y
INTERNATIONAL						
World Health Organisation[37]	Non-communicable disease	Y	Y	N	N	N
Council of the European Union[48, 50]	Cancer	Y	Y	N	N	N
RECOMMENDATIONS						
Andermann 08[6]	Genetic screening	NA	Y	N	N	N
Andermann 09[8]	Genetic screening	NA	Y	N	N	N
Andermann 10[8]	Genetic screening	NA	Y	N	N	N
Blancquaert 08[7]	Genetic screening	NA	Y	N	Y	Y
Harris 11[54]	General	NA	NU	N	Y	N
Holland 06[47]	General	NA	Y	N	N	N

Y=used or recommended N= No Information NU=Not used R=Recommended but processes unclear

In Switzerland, officials were considering the formation of a national screening body in 2010 but the website has not been updated since.[31] In Japan and the Czech Republic, we only found information for professional medical societies. The processes indicated for Belgium are based only on their recommendation for GBS and it is not clear whether the same

processes are used for other conditions. Similarly, in Australia and Italy, the screening organisations found had only covered cancer programmes.

Systems for appraising general screening programmes

All countries, with the exception of Spain, make recommendations to screen at the national level. In Spain, the regions undertake the assessment for the introduction of a screening programme and submit the review to the Ministry of Health, Social Services and Equality for reviewing and approval. Across the countries, decision-making and implementation for screening is not always made at the national level (see table 2). Belgium, France, Germany, the Netherlands, New Zealand, and the UK apply the national screening recommendations to the entire country. Of these countries, national screening organisations only in the Netherlands, New Zealand and the UK have the responsibility of implementing the screening programmes, whereas in Belgium, France and Germany implementation is delegated to regional and local authorities. In Australia, Canada, and Sweden both the decision to screen and its implementation are devolved to lower levels. The Screening Subcommittee in Australia was previously a permanent body, however it is now formed on an ad hoc basis when recommendations are to be made or changed. In France and Germany implementation is through health insurance coverage.

In Denmark, Finland, and Italy, screening bodies enforce some screening programmes and not others therefore lower level authorities have to introduce some programmes. Regional and municipal authorities in these countries are able to arrange other programmes not set out in the national screening recommendations. In the USA, the US Preventive Services Task Force (USPSTF) makes national recommendations but in previous years decision-making was left to the regional level and health insurance plans. However, as a result of the Affordable Health Care Act, screening recommendations that are graded A and B (see appendix) now have to be covered by health insurance plans nationally. Implementation of recommendations is left to the regional level.

In many countries, besides the organisation responsible for screening recommendations, other professional bodies also issue policy recommendations and statements for screening. These are usually medical specialist bodies, and in the case of GBS, they are obstetrics and gynaecology organisations. The exception is for the USA, where a public health body, the Center for Disease Control and Prevention (CDC) also develops recommendations for screening.

Topic proposal and selection method

In Canada,[44] Sweden,[55] and the USA,[39] in addition to members of the screening organisation and immediate stakeholders, anyone from the public can propose a recommendation for the screening organisation to consider. In the UK, it is any organisation from the public that is identified as a stakeholder.[4] In France, it has to be an organisation either from the government or responsible for health policy, a patient or user association, a learned medical society or university, or an expert medical professional union.[35] In Germany, the Federal Health Ministry and the Association of Insurance Doctors make suggestions.[40,41] In Belgium,

Table 2: Screening systems across countries and international bodies

Country	Screening Organisation	Recommendation	Decision	Implementation	Other organisations making screening recommendations for GBS
INTERNATIONAL ORGANISATIONS					
Europe[50]	Council of European Union	International			
World Health Organisation[37]	World Health Organisation Consultation Group	International			
COUNTRIES					
Australia[2, 23, 24, 26, 27, 52]	The Screening Subcommittee, Australian Population Health Development Principal Committee	National	Regional	Regional	Royal Australian and New Zealand College of Obstetricians and Gynaecologists
Belgium[12, 15, 22]	Superior Health Council	National	National	Local	
Canada[2, 13, 21, 44, 45]	Canadian Task Force on Preventative Health Care	National	Regional	Regional	Canadian Strategy for Cancer Control: Screening Working Group Society of Obstetricians and Gynaecologists of Canada (SOGC) Canadian Paediatric Society
Denmark[5, 48]	National Board of Health	National/ Regional	National/ Regional	Regional	
Finland[38]	National Screening Committee, Ministry of Health and Social Affairs	National/ Municipal	National/ Municipal	Municipal	
France[5, 19, 35]	Haute Autorité de Santé	National	National	Local through health insurance	
Germany[5, 41]	The Federal Joint Committee	National	National	Regional through health insurance	Working Group of the scientific medical professional societies
Italy[5, 14, 36]	National Observatory Screening	National	National/ Regional	Regional	GBS Prevention Working Group of Emilia-Romagna.
Netherlands[25, 32, 33, 49]	The Health Council National Institute for Public Health and the Environment	National	National	National	Dutch Organisation of Obstetrics and Gynaecology Dutch Association of Pediatrics (NVK)
New Zealand[2, 18, 24, 34, 53]	National Screening Advisory Committee National Screening Unit	National	National	National	Royal Australian and New Zealand College of Obstetricians and Gynaecologists
Spain[5, 29, 42]	Ministry of Health, Social Services and Equality	Regional	Regional	Regional	Spanish society of obstetricians and gynaecology
Sweden[5, 49, 55]	The National Board of Health and Welfare	National	Local	Local	
United Kingdom[4, 49]	United Kingdom National Screening Committee	National	National	National	Royal College of Obstetricians and Gynaecologists
United States of America[2, 16, 20, 30, 39, 46, 54]	United States Preventive Services Task Force	National	National coverage for a set of recommendations	Regional through health insurance	Centers for Disease Control and Prevention American College of Obstetricians and Gynaecologists American Academy of Paediatrics
Switzerland[31, 43]	Considering national recommendation and decision-making organisation				Swiss society of neonatology
Czech Republic[17]					Czech Gynaecological and Obstetrical Society
Japan[28]					Japan Society of Obstetrics and Gynaecology Japan Association of Obstetricians and Gynaecologists

for GBS, the proposal was bought to the Superior Health Council by obstetricians and gynaecologists.[12,15,22] In the Netherlands and New Zealand, the Ministry of Health/Health Council draws up the selection of screening programmes in house after conferring with various departments and groups. In the Netherlands, the parliament can also submit a

request for advice to the Council.[32,33] In New Zealand, the Ministry of health takes advice from the National Screening Advisory Committee and the sector.[34,53] In Sweden and the UK, proposals are only accepted if an attempt is made to assess the condition against the criteria used nationally.

In the UK, topics will generally be reviewed if an attempt to assess the condition against the screening criteria has been made.[4] In Sweden, the selection is based on urgency, need, current practice of condition, and programmes meeting the first 3 screening criteria (important health problem; known natural history; detectable latent stage of disease).[55] Canada[44] and the USA[39] use a similar procedure to select topics, whereby the Task Forces appoint a prioritisation working group who select a list of topics to work on during the year according to certain considerations. In Canada, a Delphi process is first used to shortlist 20-30 topics according to the following considerations – timing of most recent review, availability of new evidence, and input from primary care practitioners. To reach the final selection of 10 topics, members of the group subjectively rank the shortlist using the following considerations and mean rankings are calculated accordingly – disease burden, potential impact of recommendations in clinical practice, interest of the public or care providers, variation in care and potential for preventive service to decrease that variation, sufficiency of evidence, and new evidence, especially high-quality evidence in a stable field. The Task Force as a whole then approves the list. In the USA, topic prioritisation by the group is made according to the following considerations - public health importance, potential change to a prior recommendation if there was one, and potential for Task Force impact. In both countries, the prioritisation process takes into account the requirement that the topics for each year should cover various disease types, populations, and types of services.

Criteria used for evidence assessment

Table 3 describes every criterion in use, followed by the countries and the international health organisation that use each, for appraising general screening in the first section of columns, and the research articles that identify or recommend using them for general screening in the second section of columns. (The third section shows the criteria items recommended or used for genetic screening – see next chapter). Of the countries found in the search, no information was found for the criteria used in Belgium, the Czech Republic, Japan and Switzerland. The information found in Italy, Germany, and Finland is based on guidelines for features and assessments required for screening programmes. The Council of the European Union uses Wilson and Jungner's criteria,[3] as well as the Council of Europe recommendations,[50] so the recommendations from the Council of Europe were synthesised.[48] Similarly, the criteria reviewed by Andermann et al. (2008)[6] are Wilson and Jungner criteria[3] in addition to 54 lists of criteria found in their search. Canada[44] and the USA[39] do not use criteria and instead, a working group of experts appointed by the Task Forces develop an analytic framework and associated key questions, and contextual questions to define the scope and focus of the policy review. Through the analytical framework and key questions, the working group specify the population, intervention, and important intermediate and health outcomes for the screening programme under

consideration. Contextual questions are not part of the analytical framework but responses to these questions relate to the implementation and context of the programme. For the purpose of this review, the information for the USA and Canada in table 3 involves the key questions, commonly used contextual questions, and the review template, that direct the assessment of the screening programme. These questions will, however, differ from one topic to another and additional questions and sub-questions may be added. Finally, Harris et al. (2011)[54] provide information on an evidence process used by the USPSTF where the magnitude of benefits, harms, and net benefit is calculated. Therefore, the areas of calculating these magnitudes have been included in the table.

Frequently used criteria

There was a wide range of 46 unique items used within the screening criteria across the different countries, international health organisations and research recommendations (this does not include items found in the genetic screening criteria from the next chapter). The condition section had 7 items, the test section had 6 items, the treatment section had 6 items, and the programme section had the largest, with 27 items. The most frequently used items in the condition section were 'important health condition' (13 uses), natural history understood' (17 uses) and 'a detectable latent stage' (16 uses). In the test section, the most commonly used item was test performance (17 uses), which was stated in various ways (see below) and that 'the test should be acceptable to the population' (12 uses). In the treatment section, the most frequent items were, 'evidence-based policies covering which individuals should be offered treatment' (11 uses), and 'an effective treatment' (10 uses). In the programme section, the most frequent items were 'scientific evidence of programme effectiveness' (15 uses), 'cost-effectiveness' (21 uses), 'adequate facilities, staffing and infrastructure' (14 uses), 'monitoring of a programme' (13 uses), and 'overall benefits versus harm' (16 uses).

Criteria not in use in the UK

Compared to the UK, criteria for different countries often covered the same concepts but with more explicit wording. In table 3 the phrasing used in the UK has been highlighted for each item and the extra words used in other countries included in black but not highlighted. Most important of these, other countries explicitly state the aspects of a health problem that make it important such as 'burden of condition', 'mortality and morbidity', and 'socioeconomic impact'. In the test section, regarding test performance, the UK includes a 'simple, precise, safe and validated' test. However, other countries use a 'suitable', 'appropriate', 'accurate', 'sensitive and specific', 'reliable and reproducible' and 'efficient' test with 'high predictive values', and 'useable on a large scale'.

Important differences were also found about screening and treatment effectiveness. In the treatment section, the item 'effective treatment' is expanded for genetic screening by the inclusion of 'an offer of reproductive choice' by Blancquaert et al. (2009)[7] but has not been used in any of the countries as far as was identified. In the programme section, for screening programme effectiveness, only the UK specifically requires that the scientific evidence should be from high quality randomised controlled trials (RCTs). In New Zealand and

Blancquaert et al.'s recommendation, ideally RCT data is required, in France either RCT data or international consensus is required, and the Council of Europe recommends experimental studies. Similarly, the UK states that 'where screening is aimed solely at providing information to allow the person being screened to make an "informed choice", there must be evidence from high quality trials that the test accurately measures risk. Additional wording for screening effectiveness included 'applicable to the implementation context' and that 'other cost effective interventions implemented should be justified based on the expected impact, and values and social priorities' and ' to encourage high coverage'.

While the UK states that the whole screening programme should be ethically acceptable, other countries specifically mention the ethical issues that should be covered. Australia, France, and the Netherlands, specifically state that the screening programme should promote human rights, including 'equity', 'autonomy', and 'confidentiality'. Denmark specifically expands this to cover 'an evaluation of ethical and psychological consequences of false positive and false negative results', 'overdiagnosis', and 'stigmatisation', and the USA, and Canada, specifically consider 'the frequency and severity of the harms of work-up/screening and treatment'. With respect to informed consent, the Council of Europe and Andermann et al. (2009, 2010)[8,9] recommend that 'if screening is provided for research purposes, separate consent should be sought from patients'. Related to genetic conditions, considerations included 'neonatal screening only being justified if screening is of direct health benefit to the child' (Council of Europe); 'offering individual risk counselling' (Andermann et al., 2009; 2010)[8,9]; and expanding the item on the acceptability of mutation screening from other family members to include 'a management strategy' for other family members (Andermann et al., 2009, 2010;[8,9] Blancquaert et al., 2009[7]). Finally, ethical issues often arise around people with 'intermediate', 'indeterminate' and 'sub-clinical' or 'carrier status' results from screening. Canada and the USA have criteria about knowing 'the strength of association between intermediate and clinically relevant outcomes' and the 'presence of pseudo disease', and Australia require 'a policy for the management of high-risk individuals' in addition to those with 'intermediate' and 'negative' results.

Additional considerations were around planning and implementation issues. Australia, Canada, and the Netherlands, mention 'a recognised need', 'objectives', and a 'specified target population', as well as the 'inclusion of patient values and preferences'. The criteria item on the availability of adequate resources in the UK includes the mention of 'staffing and facilities' however other criteria also include 'infrastructure', an 'optimal screening interval', 'databases for population registers and data collection', 'training', 'education', and 'raising awareness'.

Two interesting criteria were found regarding quality assurance and monitoring – Italy mentions that 'activity of early diagnosis done outside of organised screening programs must be subjected to a quality control enabling the assessment of the adequacy and results' and the Council of Europe recommends that 'if quality assurance standards are not met in the long term it should be possible for the screening programme to be corrected, and, if this is not possible, stopped'. In the monitoring item, other countries also mention 'evaluating', which the UK does not.

Criteria unique to the UK

We found some criteria used in the UK but not used by many other countries. In the test category, we found that 'the distribution of test values in the target population should be known and suitable cut-off level defined and agreed' was only used by Australia. In the programme category, 'clinical management of the condition and patient outcomes should be optimised in all health care providers prior to participation in a screening programme' is only used by Spain. Only the UK states 'public pressure for widening the eligibility criteria for reducing the screening interval, and for increasing the sensitivity of the testing process, should be anticipated… scientifically justifiable to the public.' Regarding cost-effectiveness, only France along with the UK state 'all other options for managing the condition should have been considered to ensure that no more cost effective intervention could be introduced or current interventions increased within the resources available', although it is also recommended by Blancquaert et al. (2009)[7]. Similarly, only Spain and the UK state that 'all the cost-effective primary prevention interventions implemented as far as practicable' while Andermann et al. (2008, 2009, 2010)[6,8,9] recommend it.

Genetic mutation considerations only used in the UK state 'if the carriers of a mutation are identified as a result of screening, natural history of people with status should be understood, including the psychological implications' and 'if screening is for a mutation the programme should be acceptable to people identified as carriers and to other family members'. Only Spain along with the UK assess the criterion, 'if the test is for mutations the criteria used to select the subset of mutations to be covered by screening, if all possible mutations are not tested, should be clearly set out'.

Categorisation of criteria

As opposed to the condition, test, treatment and programme categorisation, Blancquaert et al. (2009)[7] separate criteria into 3 different decision nodes to distinctly represent the viewpoint for individuals, families, and society as a whole. At decision node A, the nature of the disease and the capacity for early detection and intervention is assessed for individuals and families. At decision node B, the relevance, efficiency and effectiveness of implementing a programme for a given target population is determined. Finally, from the societal perspective, at decision node C, the governance structures and the allocation of resources is focused on comparing the programme to other priorities and constraints. Structuring the criteria in this manner enables scientific and contextual evidence to be synthesised in a step-wise manner, making the viewpoints clear at each of these steps.

Evidence synthesis methodologies

In some countries, once a topic is identified for appraisal, it is first assessed for the type of review that is needed. For example, in the UK, the information specialist in the NSC conducts a knowledge update of the research. Based on the knowledge update, the NSC director decides whether a full external systematic review is required or not.[4] Similarly, in the USA[39] and Canada[44], once a topic is selected, a working group decides whether the topic requires a full systematic review (of every key question in the analytical framework), a target systematic review (for selected questions in the analytical framework) which is commonly

used for updates of topic reviews, or a staged review (for selected questions in the analytical framework to be answered in a ordered fashion before a full review is conducted).

In each country, the full evidence synthesis review is given to an independent body to conduct (table 4). The overall method used by each country is either a systematic review (SR) or a health technology assessment (HTA). Practice in France[35] and suggested by Harris et al. (2011)[54] is to apply modelling methods in addition to systematic reviews where possible. All countries with available information identify and include stakeholders in the process of evidence synthesis and decision-making. In many countries the stakeholders help in the initial stages of the policy review and in the end to review the evidence. However, in Canada[44], the USA[39] (as mentioned earlier), and Denmark[51], stakeholders play a larger role as they help to formulate the framework and questions that the evidence review should address. In Denmark, as well as the criteria, the stakeholder project team formulates relevant questions within the dimensions of the Danish HTA template, which include the aspects of technology, patient perspective, organisation, and economics, as well as the incorporation of ethics. In all other countries, the framework followed by the reviewers is the national screening criteria. Evidence reviewers therefore assess the evidence for the programme against the criteria.

With respect to research literature, researchers have developed guidelines on the types of evidence required to assess whether benefits outweigh harms (see appendix for example of guides). Blancquaert et al. (2009)[7] propose different types of scientific and contextual evidence to assess whether each criterion, and the decision node overall, is fulfilled. For example, in decision node A, for the criterion, “The target condition should be a common health problem with serious consequences,” the indicated supporting evidence is the availability of epidemiological data for the population and for high-risk subgroups receiving standard clinical care, establishing 1) the incidence and prevalence, 2) the average or median age of symptom onset, 3) the mortality, morbidity, quality of life and disability, 4) the utilisation rate of health care services and 5) the impact of the disease on the capacity for social adjustment or integration. Harris et al. (2011)[54] have similarly highlighted the minimal evidence required to calculate each benefit and harm. For example, to calculate the probability of an adverse health outcome in the population without screening, the minimal evidence required is 1) minimally selected population from a broad setting, reasonably representative of the population to be screened, 2) a cohort study with a reasonable measure of the health outcome over time, and 3) well-conducted retrospective studies. As far as we identified, none of the countries have implemented or developed guidelines for supporting evidence.

To assess the quality of evidence in a review, Canada[44,45], France[35], Sweden[55], and the USA[39] utilise specific tools. France provides critical appraisal questions and recommends tools for assessing specific types of study within their criteria for screening tests, programme effectiveness and economic assessment. Canada and Sweden use the methods of the Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group[58] where the quality of evidence is assessed for the combination of studies for each important outcome in question and then across the outcomes to give a score to the body of

evidence for the whole programme. The USA[39] uses similar methods developed by the USPSTF, thus the quality of individual studies is graded then combined to grade the quality for each key question, and then the quality of the key questions are combined to grade the quality of the programme as a whole. Blancquaert et al. (2009)[7] similarly recommend this type of quality appraisal following from individual studies to criteria, to overall decision node.

In the USA,[39] the examination of the quality of individual studies is based on internal and external validity. Studies are rated as good, fair, and poor. Studies with "poor" internal or external validity are not admissible for further consideration. To deliberate the level of evidence that the studies provide together for the key outcomes to each of the key questions, the Task Force assesses the evidence against 6 critical appraisal questions, which assess the research design, internal validity, external validity, number of studies and sample sizes, consistency of results, and any additional factors. In the GRADE system[58] used by Canada and Sweden, the classification of quality of evidence for each outcome is judged as high, moderate, low, or very low. If the evidence of an effect is unlikely to be influenced by additional research, the quality grade is high, if it is likely, the grade is moderate, if it is very likely the grade is low, and if it is very uncertain it is graded as very low. The quality of evidence for an individual outcome can be affected by a number of factors, such as study design, risk of bias, inconsistency across studies, indirectness of evidence, imprecision of evidence, publication bias, large effect sizes, plausible confounding, and dose-response effect.

When collating the quality of evidence for the recommendation overall, the Canadian Task Force only considers outcomes defined as critical. As recommended by the GRADE working group, if there is a critical outcome with low quality, the rating for quality of evidence across all outcomes must reflect this. Therefore, if one of the critical outcomes has a low level of evidence, the rating of evidence quality across all outcomes would be low. In the USA, the bodies of evidence are assessed for the certainty they provide for accurately estimating the expected magnitude of benefits, harms, and net benefit, as a result of a screening programme. This certainty is not a quantitative calculation, but a judgment based on the 6 critical appraisal questions used to determine the evidence across the outcomes. It is basically the "likelihood that the USPSTF assessment of the net benefit of [screening] is correct" and is graded high, moderate, or low.[39]

In the USA and Canada, the main results of the evidence review are reported in a summary of findings or outcomes table. In Canada, the table is conducted in accordance with GRADE guidelines. The Canadian Task Force also calculates the Number Needed to Screen (NNS) using the relative risk method.[44] In the USA[39] (also discussed in the literature by Harris et al)[54] the Task Force calculates the magnitude of net benefit of a screening programme. To do this, the magnitude of harms is subtracted from the magnitude of benefit. Harris et al. (2011)[54] have indicated the information required to calculate the magnitude of each – for benefit the most important factors are the probability of the adverse health outcome in the population with no screening, the degree to which the 'predictor of poor health' (identified by screening) would identify all people who would suffer the adverse health outcome, and the incremental benefit of earlier versus later treatment resulting from earlier detection.

The most important factors on the harm side of the equation are the frequency of false positives determined by the prevalence of the 'predictor of poor health' and the specificity of the screening test, the experience of people with false positives, the frequency of overdiagnosis, the experience of overdiagnosed people, and the intrinsic harms in the testing and treatment of the condition.

The two complimentary approaches that Harris et al. identify to weigh benefits and harms are outcomes tables and single metric analyses such as quality adjusted life years (QALYs). The USPSTF uses NNS and Harris et al. propose comparing the number needed to screen for benefit (NNSb) and number needed to screen for harm (NNSh). The net benefit is graded as substantial, moderate, small, and zero/negative. Thus, "substantial" net benefit indicates that benefits substantially outweigh harms, whereas "zero/negative" net benefit indicates that harms equal or outweigh benefits. There is no threshold for drawing conclusions, however in general, a substantial net benefit is considered by the USPSTF as "a large proportion of total burden of suffering from the target condition (minus the additional burden caused by the preventive service) that would be relieved from society by implementing the preventive service, even if the target condition is rare, is large" OR "a large amount of the burden of suffering would be relieved from society (minus the amount of the additional burden caused by the preventive service) by implementing the preventive service"[39]. Harris et al. have suggested using other screening programmes as a benchmark. The last step in this process is to calculate the financial costs and resources required for the screening programme (through cost-effectiveness analyses or outcome tables) and to consider whether the net benefit justifies the resources required to implement the programme.

Many countries have the evidence reports reviewed externally for quality assurance. While the UK allows consultation from anyone in the public – individuals or organisations, other countries also consult with partner organisations (Canada,[44] USA[39]), peer reviewers and experts in the field (Canada[44], Denmark[51], France[35], the Netherlands[33], New Zealand,[34,53] the USA[39]), public health bodies (Sweden[55] and Spain[42]), and stakeholders (Denmark[51]).

Decision-making processes

In order for the screening organisations to reach a decision for a screening recommendation, all countries hold a meeting where members address the key issues and reach a consensus. In New Zealand, Sweden, and the UK, it is unclear whether a structured procedure is applied. In Canada[44], the USA[39], and Germany,[41] a vote is taken, and in Canada[44] and the USA[39], a quorum is set for the required level of agreement – in Canada this is two-thirds of the members and in the USA it is 10 out of 16 votes.

Table 4: Methodology for evidence review and synthesis

Name	Inclusion of stakeholders	Review type	Framework	Who conducts review	Quality appraisal of evidence	Quality assurance of review and recommendation
COUNTRIES						
Canada[44,45]	Yes	SR	The working group develops analytic framework and key questions, approved by task force and partners	Evidence Review and Synthesis Centre	GRADE	Partner organisations, and individual peer reviewers at various stages
Denmark [51]	Yes	HTA	Criteria/Questions formulated by project team within the dimensions of the Danish HTA template	Danish Centre for Health Technology Assessment and organised project teams	Various recommendations from the Danish Secretariat of Clinical Guidelines	Reference group of stakeholders and peer reviewers
Finland[38]	N	HTA	Criteria in screening decree	Finnish Office for Health Technology Assessment (Finohta)	N	N
France[35]	Yes	SR and modelling	Criteria	Haute Autorité de Santé	Quality and validity questions provided for relevant sections	Working group Peer reviewers including patient associations
Netherlands [33]	Yes	Health assessment report	Criteria	Health council professional secretariat	N	Standing committee' of experts
New Zealand[34]	Yes	Literature review	Criteria	National screening advisory committee	N	External peer reviewers
Spain[42]	Yes	SR/HTA	Criteria	Independent expert group in collaboration with Health Technology Assessment agencies	N	Ministry of Health in collaboration with Autonomous Communicities
Sweden[55]	Yes	SR/HTA	Criteria	Expert working group appointed by the Board, commonly Swedish Council on Technology Assessment in Health Care (SBU)	GRADE	National screening council Health authority representatives Public
United Kingdom[4]	Yes	SR	Criteria	Various national experts or academic institutions	Left to reviewer	Public
United States of America[39]	Yes	SR	EPC, AHRQ staff and Task Force develop analytic framework and key questions.	Evidence-based Practice Centers	6 critical appraisal questions on internal and external validity to assess Individual study, key questions, and entire preventative service levels	Experts, federal and primary care partners, peer reviewers
RESEARCH RECOMMENDATIONS						
Blancquaert '08[7]	Recommend-ed	SR recommen-ded	Decision guide - supporting evidence required for each criterion	N/A	Recommended for every study, criterion then overall decision node	N
Harris '11[54]	Recommed-ed	SR and modelling	Search evidence to estimate health outcomes, health benefits, harms, and resources in an outcomes table	N/A	N	N

N=No information R=Recommended but processes unclear N/A=Not applicable SR=Systematic Review HTA=Health Technology Assessment

The recommendations in Canada[44], France,[35] Sweden[55], and the USA[39] are graded in line with their strength of evidence. For Sweden and Canada who follow GRADE, recommendations are graded either weak or strong. A recommendation is rated as strong if the benefit of the intervention outweighs its harms or vice versa. A recommendation is rated as weak if the benefits of the intervention probably outweigh its harms or the harms probably outweigh

the benefits. When determining the strength of a recommendation, the Canadian Task Force considers the baseline risk of the outcome, the effect size of the intervention and the precision of the effect. The quality of the evidence, the costs, patients' values and preferences, and the balance between benefits and harms are also considered.

In the USA, the strength of a recommendation is given by combining the certainty of net benefit and the magnitude of net benefit (see appendix). For example, Grade A indicates that the certainty of evidence is high that the magnitude of net benefits is substantial whereas Grade I indicates that the evidence is insufficient to determine the relationship between benefits and harms. The USPSTF also provide standardised language for their grades (see appendix).

The guide developed by Blancquaert et al. (2009)[7] provides support for decision-making processes. The separation of the criteria into the three decision nodes provides the opportunity to make the perspectives explicit in the discussion and decision of whether benefits outweigh harms from each perspective, settling the potential tensions between them. This process enhances the transparency of the evaluation of a screening programme as the evidence for each perspective is recorded and the trade-offs and tensions are made clear, thereby providing the underlying reasons for decisions and ensuring a proper balance between the perspectives.

Table 3: Comparison of criteria across general and genetic screening

No	Criteria	W&J	Aus[52]	Can[44a]	Den[48]	Fin[38]	Fra[35]	Ger[41]	Ita[36]	Neth[33]	NZ[34]	Spa[42]	Swe[55]	UK[4]	USA[39a]	WHO[37]	EU[48]	Research articles	And 08[6]	And 09/10[8,9]	Blanc[7]	Fowler[34]	Gray[54]	Harris[54]	Holland[47]	Genetic screening search	HCN[b]	Goel[b]	ESHG[b]	EC ELSI[b]	PHGF[b]	Bonham[b]
1	Suitable or well-defined candidate for screening			X		X					X				X	X				X				X								
2	The condition should be an important health problem "The overall burden of disease due to genetic conditions should merit a public health response that examines the possibility of prevention. EURORDIS advocates that this should be a global approach- rather than a piecemeal policy for each disorder separately."(PHGF 2010)	X	X		X				X	X		X	X	X	X		X		X					X							
	a. Burden of condition – Incidence & prevalence			X		X	X								X	X																
	b. Mortality and morbidity						X										X															
	c. Socioeconomic impact						X										X															
3	The epidemiology and natural history of the condition, including development from latent to declared disease, should be adequately understood/ Probability of an adverse health outcome without screening The level of risk assessed through penetrance is important. Natural course must be understood, including evolution from susceptibility to early disease. (Goel 2001) "The rarity of inherited metabolic conditions means that the methods in standard epidemiological research, which rely on a population based assessment of disease comparing populations with and without disease, will be unlikely to provide sufficient statistical power to provide classical evidence on incidence, causation, risk factors and natural history. The question at issue thus becomes: do we understand the underlying pathology and expected natural history well enough to recommend treatment that we believe to be beneficial?" (PHGF 2010)	X	X	X	X		X			X		X	X	X	X	X	X		X	X	X			X	X		X	X			X	
	a. Known strength of association between intermediate outcomes and clinically relevant outcomes for condition			X											X																	
	b. Is pseudo disease present in the apparently diseased population														X																	
4	There should be a detectable risk factor, disease marker, latent period or early symptomatic stage or increased level of genetic risk	X	X	X	X			X	X	X		X	X	X	X		X		X	X	X				X			X			X	
5	All the cost-effective primary prevention interventions implemented as far as practicable.						X					X		X					X	X									X			
6	If the carriers of a mutation are identified as a result of screening, natural history of people with status should be understood, including the psychological implications.													X																		
7	Current clinical practice			X											X																	

a. Do not use criteria b. Only assessed against NSC criteria and not against other criteria found in the course of extracting data

Abbreviations: W&J=Wilson & Jungner, Aus=Australia, Can=Canada, Den=Denmark, Fin=Finland, Fra=France, Ger=Germany, Ita=Italy, Neth=Netherlands, NZ=New Zealand, Spa=Spain, Swe=Sweden, UK=United Kingdom, USA=United States of America, WHO=World Health Organisation consultation group, EU=Council of Europe, And=Anderman er al., Blanc=Blanquaert et al., HCN= Health Council of Netherlands, ESHG= European Society of Human Genetics, EC ELSI= European Commission Ethical Legal and Social Implications of genetic testing, PHGF=PHG Foundation

- Criteria wording used by UK NSC
- Criteria wording used in other general screening criteria
- Criteria wording used in other genetic screening criteria
- Additional criteria suggested for genetic screening programmes

Table 3: Comparison of criteria across general and genetic screening

No	Criteria	W&J	Aus	Can[a]	Den	Fin	Fra	Ger	Ita	Neth	NZ	Spa	Swe	UK	USA[a]	WHO	EU	Research articles	And 08	And 09/10	Blanc	Fowler	Gray	Harris	Holland	Genetic screening search	HCN[b]	Goel[b]	ESHG[b]	EC ELSI[b]	PHGF[b]	Bonham[b]
	Test					X																										
8	Suitable or appropriate test or examination	X		X	X			X	X		X		X				X							X	X		X	X	X	X		
	a. Test should be simple						X					X		X		X									X							
	b. Test should be precise													X																		
	c. Test should be accurate			X											X										X							
	d. Test should be sensitive and specific (including true positives, false positive, overdiagnosis)		X	X											X		X							X	X		X					
	e. Predictive values of test should be evaluated/relatively high		X		X												X							X			X			X		
	f. Test should be valid		X		X		X					X		X							X											
	g. Test should be safe		X									X		X		X					X				X							
	h. Test should be reliable and reproducible						X			X		X													X							
	i. Test should be efficient				X							X																				
	j. Usable on a large scale																				X											
9	The test and [further investigation] should be acceptable to the population including sub groups such as target participants who are from culturally and linguistically diverse backgrounds, disadvantaged groups, and disabled	X	X		X		X			X		X	X	X			X		X	X					X			X				
10	The distribution of test values in the target population should be known and suitable cut-off level defined and agreed on whom to categorize as "screen positive" "screen negative" and "screen indeterminate"		X											X		X				X					X							
11	There should be agreed evidence based policy for each group following disclosure of screening results, regardless of rurality, ethnicity, socioeconomic status, or disadvantage status:																															
	a. On further diagnostic investigation assessment and support of individuals with a positive test result and on the choices available to those individuals		X		X		X					X		X		X	X			X		X					X	X		X		
	b. Negative screening tests, including providing information		X																	X												
	c. Intermediate results																			X												
12	If the test is for mutations the criteria used to select the subset of mutations to be covered by screening, if all possible mutations are not is tested, should be clearly set out.											X		X																		
13	New technologies for screening/and or intervention														X																	
14	Systematic case finding followed by systematic cascade testing is intermediate between population screening and testing of high-risk individuals and should also be considered, according to the same criteria as population genetic screening ESHG PPPC (2003)																												X			
15	Rapid advances in technology may make it possible to screen large number of disorders or traits simultaneously. It will then become difficult if not impossible to provide proper information about each of the conditions and traits screened. Our recommendation is to authorise packages only when there is enough consistency in the characteristics of the conditions screened to allow properly informed consent from the consumer. ESHG PPPC (2003)																												X			

Table 3: Comparison of criteria across general and genetic screening

No	Criteria	W&J	Aus	Can[a]	Den	Fin	Fra	Ger	Ita	Neth	NZ	Spa	Swe	UK	USA[a]	WHO	EU	Research articles	And 08	And 09/10	Blanc	Fowler	Gray	Harris	Holland	Genetic screening search	HCN[b]	Goel[b]	ESHG[b]	EC ELSI[b]	PHGF[b]	Bonham[b]
	Treatment					X																										
16	There should be an effective treatment or intervention to benefit premature mortality, benefit quality of life, or alter the course of disease for patients identified through early detection		X	X			X	X			X			X		X	X				X				X			X	X			
	a. Evidence of early treatment leading to better outcomes than late treatment						X		X			X	X	X	X		X							X				X				
	b. Treatments available that make a difference in intermediate outcomes when the disease is caught early, or detected by screening														X																	
	c. And for genetics, the intervention should offer a reproductive choice based on an improved risk assessment																				X											
17	There should be an accepted treatment	X	X		X					X			X				X		X	X						X						
18	The treatment must be accessible		X								X						X															
19	The treatment must be available		X														X								X							
20	There should be agreed evidence based policies and referral systems covering which individuals should be offered treatment and the appropriate treatment to be offered or the management of any abnormalities	X	X		X		X			X		X		X		X	X		X			X									
	a. There needs to be an established policy/special considerations for the management of individuals who are identified at high risk of developing the disease or condition		X																													
21	Clinical management of the condition and patient outcomes should be optimised in all health care providers prior to participation in a screening programme "The introduction of screening often provides the necessary stimulus to agree national professional treatment guidelines and frequently leads to the improved organisation of services for patients detected" (Bonham 2013)											X		X															X	X		X

Table 3: Comparison of criteria across general and genetic screening

No	Criteria	W&J	Aus	Can[a]	Den	Fin	Fra	Ger	Ita	Neth	NZ	Spa	Swe	UK	USA[a]	WHO	EU	Research articles	And 08	And 09/10	Blanc	Fowler	Gray	Harris	Holland	Genetic screening search	HCN[b]	Goel[b]	ESHG[b]	EC ELSI[b]	PHGF[b]	Bonham[b]
	Programme																															
	Effectiveness					X																										
22	Have scientific evidence of screening programme effectiveness (reduces morbidity and mortality)		X	X			X			X	X	X	X		X		X		X	X	X						X		X	X		
	a. There should be evidence from high quality Randomised Controlled Trials that the screening programme is effective in reducing mortality or morbidity applicable to the implementation context. Where screening is aimed solely at providing information to allow the person being screened to make an "informed choice", there must be evidence from high quality trials that the test accurately measures risk. The information that is provided about the test and its outcome must be of value and readily understood by the individual is screened										X			X							X		X									
	b. Have evidence from experimental studies																X															
	c. Where screening is aimed solely at providing information to allow the person being screened to make an "informed choice", there must be evidence from high quality trials that the test accurately measures risk. The information must be of value and readily understood													X																		
	d. For genetic conditions, "It is impossible to obtain evidence of effectiveness of screening programmes from randomized trials because of the rarity, complexity and heterogeneity of the conditions. Not only will there be insufficient patients to generate the necessary statistical power, but also there would be significant ethical considerations in allocating patients to a non-screening group in the light of rapid advances in dietary and other aspects of management. The report discusses the 'next best' study designs (observational cohorts) in which screened and clinically detected cohorts are compared in nearby geographic areas with similar services or sequentially with groups closely related in time before and after a screening programme is put in place."(PHGF 2010) "The demand for evidence from high quality randomised clinical trials (RCT) is very difficult to satisfy. The rarity of the conditions linked to their intrinsic heterogeneity would demand multinational studies over many years to demonstrate benefit in some cases and the ethics of withholding screening to establish valid comparator groups may be difficult to justify in the face of mounting public pressure. This does not deny the need to establish good quality outcome studies linked to existing screening programmes based upon agreed case definitions and consensus approaches to treatment. Indeed, this is a pressing priority for existing programmes and could be viewed as a pre-requisite for future development. Nevertheless, the lack of RCT evidence in itself cannot be viewed as a barrier to the introduction of screening."(Bonham 2013)																														X	X
23	Cost-effectiveness: The opportunity cost of the screening programme (including testing, diagnosis and treatment, administration, training and quality assurance) should be economically balanced in relation to expenditure on medical care as a whole (i.e. value for money) applicable to the implementation context. Assessment against this criteria should have regard to evidence from cost benefit and/or cost effectiveness analyses and have regard to the effective use of available resource. "Economic criteria alone cannot be used to justify a screening programme." ESHG PPPC (2003)	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X		X		X		X		X	X		
	a. All other options for managing the condition should have been considered to ensure that no more cost effective intervention could be introduced or current interventions increased within the resources available. This should be justified based on the expected impact as well as the values and social priorities at stake.						X							X			X				X		X		X				X			
	b. Cost effective to encourage high coverage.		X																													

Table 3: Comparison of criteria across general and genetic screening

No	Criteria	W&J	Aus	Can[a]	Den	Fin	Fra	Ger	Ita	Neth	NZ	Spa	Swe	UK	USA[a]	WHO	EU	Research articles	And 08	And 09/10	Blanc	Fowler	Gray	Harris	Holland	Genetic screening search	HCN[b]	Goel[b]	ESHG[b]	EC ELSI[b]	PHGF[b]	Bon[b]
	Programme																															
	Planning and implementation																															
24	Respond to a recognised need		X							X									X	X	X											
25	Have a clear definition of the objectives of the programme, the roles and responsibilities, the expected health benefits, and the financing required from the outset		X							X							X		X	X	X											
26	Identify the target population which stands to benefit from screening		X							X		X					X		X	X	X	X		X								
27	Patient values and preferences included in screening policy-making			X																X	X												
28	Case-finding should be a continuing process and not a once and for all process with a clearly defined and optimal interval	X	X	X	X		X			X					X		X		X	X												
	a. Ages when screening should be stopped														X																	
29	Public pressure for widening the eligibility criteria for reducing the screening interval, and for increasing the sensitivity of the testing process, should be anticipated. Decisions about these parameters should be scientifically justifiable to the public.													X																		
	Organisation, Infrastructure, Workforce, and facilities																															
30	Relevant programme organisational aspects should have clarified (to achieve national equivalence)				X	X			X				X																			
	a. A detailed description of staff training/ Provision of screening programmes requires that training in techniques and interpretation of screening tests is included in undergraduate and post-graduate medical teaching programmes.				X												X					X										
	b. Awareness programmes should be organised for target population and health professionals						X										X															
	c. The programme should integrate a coherent set of education, training, practice, test, care, clinical services, and programme management		X							X									X	X	X											
31	Adequate workforce, facilities, education, organisation and infrastructure for testing, diagnosis, follow-up assessment, treatment and programme management should be available prior to the commencement of the screening programme	X	X		X		X	X		X			X	X		X	X		X	X	X	X					X	X	X			
	a. "This criteria proves difficult to satisfy, as a result of facilities for the effective diagnosis and treatment of rare disorders are not optimal even in highly developed countries. In addition, relatively few agreed international guidelines for treatment exist and there are very few agreed case definitions to guide classification."(Bonham 2013)																															X
32	Registration system - a database capable of providing a population register for people screened and health information collected for the programme		X		X				X								X															
	a. A detailed description of test result dissemination				X				X								X															
33	Feasible program within the health System					X					X	X	X				X															
34	A detailed description of steering committee (make-up, competence)				X																											
35	Involve multiple disciplines and professions								X																							
36	Programme managers should be accountable to the governance structure at the government level.																				X											
37	The governance structure at the government level should be accountable to society, in terms of both the overall performance of the programme and the implications for society																X				X											

Table 3: Comparison of criteria across general and genetic screening

No	Criteria	W&J	Aus	Cana	Den	Fin	Fra	Ger	Ita	Neth	NZ	Spa	Swe	UK	USA[a]	WHO	EU	Research Recommendations	And 08	And 09/10	Blanc	Fowler	Gray	Harris	Holland	Genetic screening search	HCN[b]	Goel[b]	ESHG[b]	EC ELSI[b]	PHGF[b]	Bon[b]
	Programme																															
	Monitoring and quality assurance																															
38	There should be a plan for managing, monitoring and evaluating the screening programme. The programme must evaluate participation, and the percentage of people screened in the target population, the technical quality of testing and the quality of diagnosis and treatment provided as a follow-up for persons with a positive test result.		X	X			X			X		X	X	X			X		X	X	X	X			X		X	X	X	X		
39	An agreed set of quality control and assurance standards to minimise potential risks of screening		X				X		X	X		X		X			X		X	X	X	X										
	a. Activity of early diagnosis done outside of organised screening programs must be subjected to a quality control enabling the assessment of the adequacy and results.								X																							
	b. A limited number of the appraisal criteria and indicators should be validated; they should be chosen at the design stage and be based on the results of the literature review or the opinion of a panel of experts						X																									
	c. If quality assurance standards are not met in the long term it should be possible for the screening programme to be corrected, and, if this is not possible, stopped.																X															
40	As the prevalence of genetic traits often varies among populations, screening programmes may be better targeted to subpopulations with high prevalence if the community agrees to have a focused health-care programme. In populations composed of subpopulations with different genetic background, the test should be selected according to population substructure. ESHG (2003)																												X			
	Acceptability and ethical issues																															
41	There should be evidence that the complete screening programme (test, diagnostic procedures, treatment/ intervention) is clinically and socially acceptable to health professionals and the public.										X	X		X					X								X	X	X			
42	There should be evidence that the complete screening programme (test, diagnostic procedures, treatment/ intervention) is ethically acceptable to health professionals and the public.					X			X		X	X	X	X					X								X					
	a. An evaluation of the ethical and psychological consequences for the examinees				X																			X								
	i. An evaluation of the consequences of "false positive" and "false negative" test results/experience of overdiagnosis				X																											
	b. An evaluation of stigmatisation				X																											
43	Promotion of human rights, including upholding the principles of autonomy and confidentiality		X				X			X							X		X	X	X											
44	Promote equity and access to screening for the entire target population, regardless of socio-cultural and economic availability		X	X			X		X	X						X	X		X	X	X											
45	Informed choice: Evidence-based information, explaining the consequences of testing, investigation and treatment, should be made available to potential participants to assist them in making an informed choice		X		X		X		X	X			X	X			X		X		X						X		X			
	a. If the programme is provided as a service and conducted also for research purposes, the decision to make available personal medical data stemming from the screening programme for research purposes should be taken freely, without undue pressure/ The consent process should be separate to consent for clinical purposes																X			X												
46	Neonatal screening can only be justified if the intervention is of direct health benefit to the child. Otherwise screening should be postponed until the child can decide for itself.																X															
47	If screening is for a mutation the programme should be acceptable to people identified as carriers and to other family members/ family members should be implicated throughout the screening process/The proposed test and intervention should be part of a coherent management strategy for all individuals in the target population who undergo screening and their families													X						X	X											
48	There should be an education programme in place from the outset of the programme and individual risk counselling should be available throughout the screening process																			X												

Table 3: Comparison of criteria across general and genetic screening

No	Criteria	W&J	Aus	Can[a]	Den	Fin	Fra	Ger	Ita	Neth	NZ	Spa	Swe	UK	USA[a]	WHO	EU	Research articles	And 08	And 09/10	Blanc	Fowler	Gray	Harris	Holland	Genetic screening search	HCN[b]	Goel[b]	ESHG[b]	EC ELSI[b]	PHGF[b]	Bon[b]
	Programme																															
	Overall benefits versus harm																															
49	The benefit from the screening programme should outweigh the physical and psychological harm (caused by the test, diagnostic procedures and treatment)		X	X			X		X	X	X	X	X	X	X		X		X	X	X		X	X				X	X			
	a. Frequency and severity of harms of workup / screening test			X											X									X								
	b. Harms of treatment														X									X								
	c. A programme should only be considered if benefits clearly outweigh harm. Screened patients, before relatives. ESHG PPPC (2003)																												X			

a. Do not use criteria b. Only assessed against NSC criteria and not against other criteria found in the course of extracting data

Abbreviations: W&J=Wilson & Jungner, Aus=Australia, Can=Canada, Den=Denmark, Fin=Finland, Fra=France, Ger=Germany, Ita=Italy, Neth=Netherlands, NZ=New Zealand, Spa=Spain, Swe=Sweden, UK=United Kingdom, USA=United States of America, WHO=World Health Organisation consultation group, EU=Council of Europe, And=Anderman er al., Blanc=Blanquaert et al., HCN= Health Council of Netherlands, ESHG= European Society of Human Genetics, EC ELSI= European Commission Ethical Legal and Social Implications of genetic testing, PHGF=PHG Foundation

- Criteria wording used by UK NSC
- Criteria wording used in other general screening criteria
- Criteria wording used in other genetic screening criteria
- Additional criteria suggested for genetic screening programmes

Systems and criteria for appraising genetic screening programmes

Cameron L, Burton H

Introduction

This chapter describes a review to identify and compare the criteria used by other countries or proposed in the literature to appraise genetic screening programmes and compare this specifically to the current NSC criteria as well as summarise the regulatory structures responsible for decision-making across countries, with particular reference to genetic screening programmes.

Methods

Search strategy

To find information on criteria used for genetic screening, one author searched Medline, Embase, ASSIA, and SSCI on Web of Science on 27 January 2014 based on the methodology set out by FS and STP in the previous chapter. In order to capture articles specifically concerned with appraisal of genetic/genomic screening, the initial search terms relating to policy and screening were combined with genetic/genomic search terms. The detailed search strategies are given in the appendix. One author (LC) reviewed titles, abstracts, and full text articles to identify papers, which focused on genetic screening programmes. Articles were excluded if they did not refer to genetic screening policy or appraisal, or did not contain sufficient information.

Reference lists from included articles were also examined for further sources, and hand searching provided additional articles. After finalising a list of included articles, HB was consulted to see if there were any significant omissions from the list. Articles were then grouped in two categories; the first describing criteria for appraising genetic screening programmes (which are included in this review), and the second discussing ethical legal and social (ELS) issues pertinent to genetic screening programmes. See PHG Foundation report[59] for detailed discussion of ELS issues and suggested modifications to NSC criteria.

Data extraction and synthesis

It was not necessary to assess the final list of included papers for quality, due to the descriptive nature of the review. One author (LC) extracted the data on customised extraction forms which included: author; title; year; country; type of article; purpose of article; the area of genetic testing; source/ participants; main findings /criteria. LC synthesised the criteria from the articles by mapping them on to the original Wilson and Jungner criteria,[3] and the current NSC criteria.[4] The criteria items from each source were not assessed against those of the other included sources, in the course of extracting the data. Sources of information on international decision-making structures, provided from the

review in the previous chapter on general screening were scrutinised for information on genetic/genomic screening appraisal and any special arrangements for undertaking decision-making in this area, particularly delegation of decision-making arrangements with regard to genetics, were also recorded.

Results

Flow of information

A total of 3582 articles were identified from the searches, and 87 abstracts or full texts were examined (see figure 2). Two of these were included in the final analysis and a further 4 were identified from websites and hand-searching.

General characteristics of information found

A total of 6 articles explicitly discussed criteria for appraising genetic screening programmes.[60-65] One report was from the Netherlands (HCN),[60] another from the USA (Goel, 2001),[65] and a further 4 were concerned with genetic screening programmes across Europe. These include the European Commission's report on the ethical, legal, and social issues for genetic testing (EC ELSI),[62] the European Society of Human Genetics (ESHG),[61] the Public Health Genomics Foundation (PHGF),[64] and Bonham (2013)[63].

Systems for appraising genetic screening programmes

The structure and function of the decision-making bodies varied amongst different countries, but some fundamental similarities with the UK situation were observed. Many countries have devolved responsibility for appraising all screening programmes (including genetic screening programmes) to a specialist body, which produces recommendations, which are then acted upon at a national or regional level. In relation to expanded newborn screening panels, Bonham (2013) notes that in countries where the 'professional genetics' community dominates decision-making, a larger range of conditions are included, as opposed to those countries where the responsibility rests mainly with the public health specialists and epidemiologists.[63]

The non-uniform approach to governance and decision-making can result in regional variation in screening, a situation that has been avoided in the UK through national implementation of programmes. In the course of our review, we found no current evidence of any genetic screening authorities acting independently from generic screening or healthcare bodies in other countries. However, some countries such as the Netherlands do have input from a specialist sub-committee on this topic. Some countries partition responsibility for different aspects of genetic screening such as cancer, prenatal screening or reproductive genetic testing.

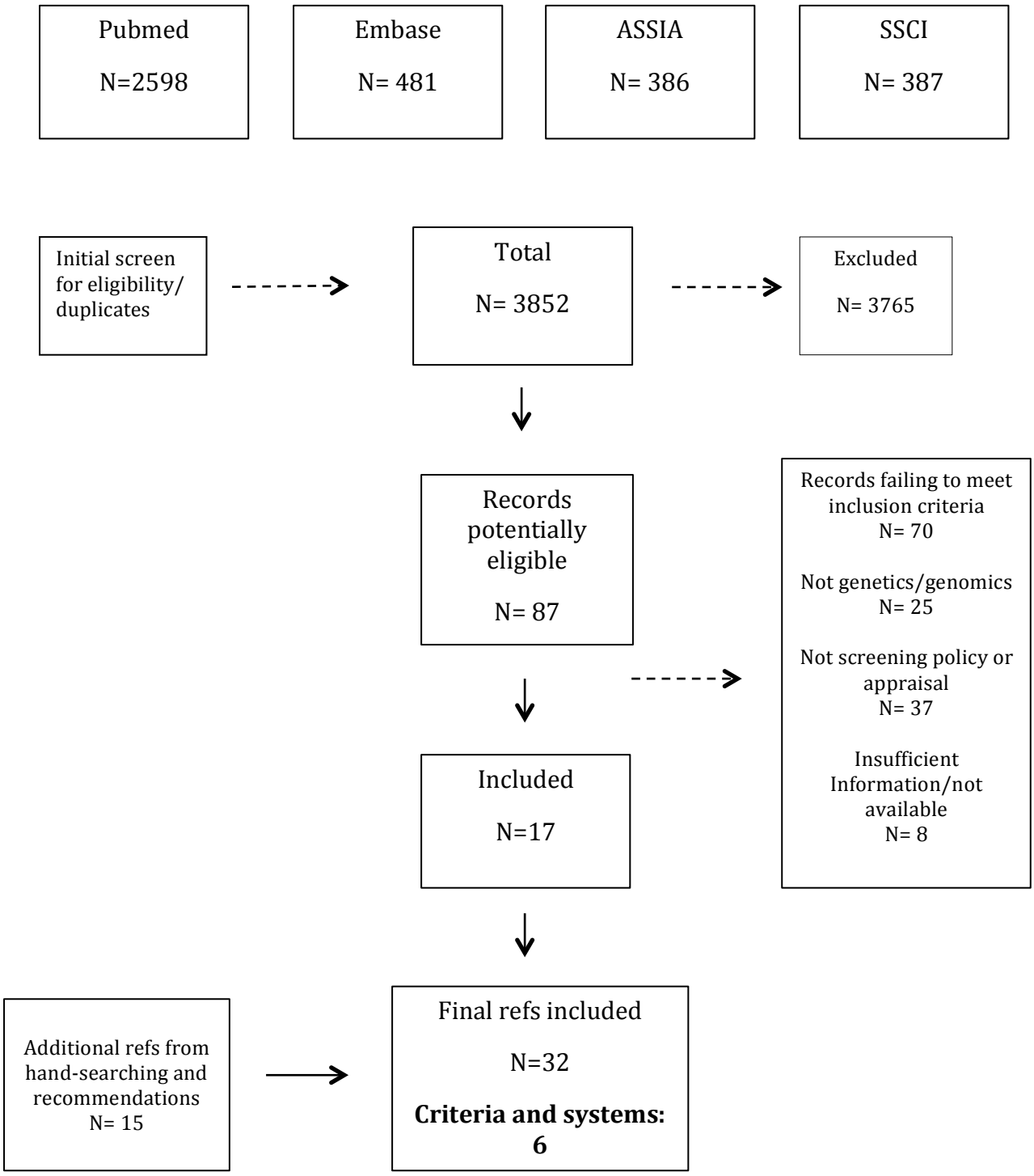

Figure 2: Flowchart of information from the genetic screening search

Other professional bodies contribute to the decision-making process by making recommendations and statements on screening, and in the case of genetic screening, such organisations tend to include genetic organisations, and other involved clinical groups such as obstetrics and gynaecology, cancer and rare disease groups (see table 5).

Table 5: Screening systems for genetic screening appraisal		
Country	**Organisation**	**Advisory genetic organisations**
COUNTRIES		
Australia	The Screening Subcommittee, Australian Population Health Development Principal Committee	Re newborn screening: A joint committee of the Human Genetics Society of Australasia and Royal College of Physicians of Australasia advises on policy, quality assurance, and other matters
Belgium	Superior Health Council	Higher Council on Human Genetics
Canada	Canadian Task Force on Preventative Health Care	Canadian Strategy for Cancer Control: Screening Working Group Society of Obstetricians and Gynaecologists of Canada (SOGC) Canadian Paediatric Society
Denmark	National Board of Health	Danish Council of Ethics, Danish Centre for Human Rights
Finland	National Screening Committee, Ministry of Health and Social Affairs	Society for Medical Genetics
France	Haute Autorité de Santé	National Ethical Consultative Committee for the Life and Health Sciences in France, Genetics and Medicine National Advisory Committee on Bioethics National College of Gynaecologists and Obstetricians
Germany	The Federal Joint Committee	The German Society of Human Genetics
Italy	National Observatory Screening	The Italian Committee on Bioethics
Japan	Ministry of Health, Labour and Welfare	
Netherlands	The Health Council National Institute for Public Health and the Environment, with committee on genetic screening	
New Zealand	Minsitry of Health, National Screening Advisory Committee National Screening Unit	
Spain	Ministry of Health, Social Services and Equality	
Sweden	The National Board of Health and Welfare	Swedish Society for Medical Genetics
Switzerland	Responsibility of Swiss Medical Board	The Swiss Academy of Medical Sciences
UK	UK National Screening Committee	
USA	U.S. Preventive Services Task Force	American Medical Association (AMA); American College of Medical Genetics (ACMG); American Society for Human Genetics (ASHG); American Academy of Pediatrics (AAP)
INTERNATIONAL BODIES		
Europe	Council of European Union	
WHO	WHO Consultation Group	Committee on Genetic Screening

Genetic screening criteria

From the included articles describing criteria, the number of discrete criteria discussed varied from three for the PHGF report[64] and the paper by Bonham[63] to 21 for the HCN[60]; this compares with the NSC's 22 criteria[4]. Some discrete criteria in the sources were encompassed within 1 criterion listed by the NSC. Table 3 under the section 'genetic screening search' shows these criteria from the international models or the proposed systems in the literature.

Criteria with proposed modifications or not explicitly stated in NSC list

Table 3 shows the additional wording used in other genetic criteria not used by the NSC in blue. Although the NSC criteria discuss the need to understand the epidemiology, a detectable risk factor, and disease marker, there is no explicit mention of the level of risk

assessed through penetrance, as described at the Canadian Crossroads Workshop by Goel (2001)[65]. In the context of the natural course of disease, Goel explicitly address the question of genetic risk and susceptibility, and advise an extension of the original evaluation framework to include the natural history of evolution from susceptibility to clinical presentation.

The criteria proposed by the ESHG Public and Professional Policy Committee (PPPC)[61] make mention of benefits and harms to screened patients before relatives, which is not explicitly mentioned by the NSC (although the NSC may consider this implicit in its assessment of benefits and harms). The ESHG recommendations also state that economic criteria alone cannot be used to justify a screening programme. The NSC approach balances the economic impact of new screening programmes with health benefits and issues such as acceptability, but does not explicitly caution against the over-emphasis of economic considerations.

The PHGF's report[64] on expanded newborn screening for inherited metabolic disorders proposes modifications to the NSC criteria which would assist in the evaluation of screening programmes for rare diseases. The proposals emphasise the need to consider the overall burden of rare diseases, rather than considering the prevalence of each condition separately. In relation to epidemiological studies, they note that, due to the rarity of these conditions, population based assessments are unlikely to provide sufficient statistical power and that decision-making with regard to screening programme implementation should be based on an understanding of the underlying pathology and expected natural history of the disease. They suggest that observational cohorts be used as evidence for the efficacy of a screening programme in place of RCT evidence, which cannot be attained due to the rarity of the conditions and ethical concerns about withholding potentially beneficial treatment.

Bonham (2013)[63] also refers to the issue of RCT data along with proposed modifications to two further NSC criteria. These criteria stipulate that clinical management of the condition and patient outcomes should be optimised along with facilities and staffing for testing and treatment before the implementation of a screening programme. Bonham argues that this is difficult to achieve in the rare disease sphere, and questions whether it is right to penalise this group of patients further by not introducing screening. He notes that, in practice, the implementation of a screening programme often provides the necessary stimulus to agree national professional treatment guidelines and frequently leads to the improved organisation of services for patients detected through screening.

Additional criteria

While the criteria relating to the condition and treatment were largely the same amongst the different sources, the greatest divergence was seen in the criteria relating to test and programme organisation. This is shown in table 3 in green. Within the test category, the ESHG PPPC criteria discuss the impact of new technologies, and caution against introducing screening because it is technologically possible, but advise waiting until there is 'enough consistency in the characteristics of the conditions screened to allow properly informed consent from the consumer.' The ESHG PPPC criteria set out some issues with regard to

working definitions: stating that systematic case finding followed by cascade testing is intermediate between population and high-risk screening. The report recommends that this type of screening should not be done ad hoc but is subject to the same decision-making processes as population screening programme proposals (focusing on benefits and harms and implementation). The ESHG PPPC also introduces the concept of heterogeneity of populations with respect to their genetic background in the consideration of genetic screening programmes. It notes that as the prevalence of genetic traits often varies among populations, it may be beneficial if the community agrees to have a focused healthcare programme, to target screening programmes within sub-populations with high prevalence. In populations composed of sub-populations with different genetic backgrounds it recommends selecting the tests according to population substructure.[61]

Criteria unique to NSC

The majority of the criteria listed by the NSC are also described in the sources on genetic screening, but similar to the general screening findings, five criteria are exclusive. Three of the unique NSC criteria focus on genetic screening programmes: (1) the needs of genetic mutation carriers, in terms of the natural history of this group and in particular the psychological implications of screening; (2) the importance of defining and explaining the subset of mutations which will be tested for if this does not cover all known mutations; and (3) highlights the importance of acceptability of the programme for carriers of a mutation and their families. A further two criteria not mentioned in other genetic criteria lists but were not specifically related to genetics were 'public pressure for widening eligibility criteria' and 'the need to understand the distribution of test values in the population, to allow determination of suitable cut-off values'.[4]

Discussion

Key findings

This research has highlighted that across the world, governments have established evidence-based policy processes in the screening arena. All countries analysed recognise the need and use for robust principles in assessing the introduction of a screening programme. To do this, they have accepted some form of the principles originally developed by Wilson and Jungner,[3] expanded the principles, and apply them most commonly as a set of criteria. In order to assess these principles, all countries are in strong agreement that a scientific evidence base must be called upon by independent experts to address the key principles and ensure that the benefit of screening does outweigh the harm. The final decision requires judgment to the principles and related discussion between members of the screening organisation. The USA[39] and Canada[44] have a very explicit and systematic approach to their decision-making, and although they do not explicitly use a criteria list, they echo Wilson and Jungner's criteria within their processes. In some countries, the screening organisations are responsible for deciding and implementing the recommendations that result from the policy-making process nationally, however in other countries governments at the regional and local level can set policies that may not necessarily have to adhere to the evidence-based recommendations. Similarly, in countries with insurance-based systems, screening is not population-based and although it is reimbursed it may differ vastly from one region and even one person to another.

Implications for the UK

The UK is quite advanced in the assessment of screening programmes and is viewed favourably in the literature as a "model" approach,[1] and having the "most integrated and evidence-based screening programmes in the world"[2]. All screening proposals suggested by stakeholders are reviewed so long as an attempt to meet criteria has been made. The review involves key stakeholders and experts conducting evidence reviews against a set of criteria with 22 well-established items that cover the major issues and which have been used by other countries. The review is shared for public consultation and a meeting is held to form a decision.[4] Comparing the UK policy-making processes to those of different countries, international bodies, and general and genetic research recommendations, we found important differences in each of the policy-making areas. With respect to topic selection, the UK only accepts programme proposals from stakeholder organisations and does not prioritise screening programmes to review. In Sweden[55], the USA[39], and Canada[44], topics can be suggested by anyone in the public and are then selected for review by using specific considerations and appointing a working group to narrow down the proposals.

The criteria not used by the UK, for the most part, are items that are covered broadly but not explicitly mentioned in the UK NSC criteria, and are not left out altogether. The criteria from other countries which are not in the UK NSC criteria include ethical issues, in particular consideration of the consequences of false positive and false negative test results, overdiagnosis, stigmatisation, the promotion of human rights, confidentiality, autonomy,

equity and access, prioritising screened patients over relatives, counselling services, and for people with indeterminate/intermediate test results, consideration of the strength of association between test results and clinically relevant outcomes, and policies for the management of such high risk individuals. There are also process and implementation considerations explicitly mentioned in criteria from other countries including a recognised need, objectives, and the target population of the programme, the inclusion of patient values and preferences in the policy-making process, consideration of the optimal screening interval, infrastructure needs, and the need for databases of population registers and data collection systems. A range of wording for test performance criteria is found across countries including accurate, sensitive, specific, high negative and positive predictive value, reliable, reproducible, efficient, and usable on a large scale. Quality assurance considerations used by other countries include quality assurance for screening outside of organised screening programmes and that if quality standards can no longer be met, then programmes should be corrected or stopped. For genetic conditions, the recommendations highlight that requirements from the current list of criteria, such as evidence from RCT data and optimised clinical management and facilities may be impossible to meet due to the rarity and heterogeneity of genetic conditions. Finally there is no consensus across countries and recommendations about whether the aim of screening is to reduce morbidity or mortality, or whether other outcomes without these could form the basis for a screening programme such as informed reproductive choice.

Additional criteria not covered by the UK NSC were found only from the genetic screening criteria lists, advising; (1) caution against introducing screening because it is technologically possible, but waiting until there is enough consistency in the characteristics of the conditions screened to allow properly informed consent from the consumer, (2) subjecting cascade screening to the same procedures as population screening, and (3) selecting genetic tests according to population substructure and considering focused screening programmes where populations are composed of sub-populations with different genetic backgrounds.[4]

The UK generally uses the same methods of systematic reviewing as most other countries for synthesising evidence. However in many countries the procedures within the evidence review are often prescribed, and therefore standardised across reviews. In the UK the review is commissioned to various national experts which means that the procedures within the review may differ. For evidence appraisal, Canada,[44,45] and Sweden,[55] use the standardised tool, GRADE,[58] and the USA have adopted their own methods very similar to GRADE, which they apply to every review.[39] The benefit of this approach is that the evidence is appraised at the higher level of key outcomes and then across all key outcomes for the programme overall as opposed to the individual study. Additionally, the strength of a recommendation if also graded, however, this might be more appropriate for countries where the decision-making responsibility is not with the screening organisations but rather devolved to the regional and local level. Nevertheless it may be helpful to link the strength of a recommendation to the timing for the update of a screening recommendation.

Although Harris et al (2011)[54] suggest calculating the magnitude of net benefits (used in the USA) instead of a criteria/checklist approach, it would be possible to use it in conjunction

with the criteria to summarise all of the criteria in one measure. After subtracting the harms from the benefits of screening, whether the net benefits would justify the resources needed could be considered in a cost-effectiveness analysis. Harris et al[54] provide a guide to carry this out with the possible approaches to do so. Finally, in the research literature we found tools that set out and standardise the type of evidence required for each criterion.[7,54] This is not something used by the UK NSC or by other countries.

Compared to the UK, other countries have a voting system to make decisions. To assist the decision-making, the decision support guide developed by Blancquaert et al. (2009)[7] provides the opportunity to make the different perspectives explicit in the discussion and decision of whether benefits outweigh harms.

Strengths, limitations, and future research

The methods of this review, emphasising the reference list search and including an expert panel, enabled systematic searching for information across different countries. With respect to the general screening review, there were two reviewers to sift, extract, and synthesise the information, and an effort was made to translate papers not in English. However, our findings must be considered in light of some limitations. Firstly, while an effort was made to informally translate papers, there may be some error in translations. Secondly, while a huge amount of work was carried out to find the correct screening organisations and processes, in some countries the screening organisations do not [yet] cover all screening programmes. For example, in Australia and Italy the associated screening bodies only cover cancer. However these countries do have antenatal and other screening programmes,[1,2] making it unclear which body is responsible for screening other conditions. It may be that there is not a single national body for reviewing screening programmes. Similarly, in Belgium only GBS was covered. Our search strategy helped to identify screening organisations and policy-making processes in countries that that use them but did not necessarily identify countries that do not have screening organisations. Finally, this search only resulted in countries from Australasia, North America, and Western Europe. As far as we are aware, Eastern and Central Europe, Asia, Africa, or Latin America do not have screening bodies.[49,66] Specifically for the genetic screening review, resource considerations meant that only one reviewer was able to sift, extract and synthesise the data.

Based on the results of information found, the policy processes in place are clear but what actually happens in practice and to what extent recommendations are adhered to is unclear. These gaps, especially on the decision-making discussions and procedures that occur in the meetings, may not be filled through a literature search. Interviews with key members from the respective countries may be more useful to obtain information on the details of policy-making practice. Secondly, while this review took a national view of policy-making processes in screening, future researchers may wish to explore the regional and local level processes that exist in countries where recommendations and decision-making responsibilities are devolved.

Conclusion

We sought to understand the policy-making processes used internationally when deciding whether or not to implement a screening programme. We have identified differences in the criteria and tools that are used in other countries compared to the UK, which can aid the UK NSC in choosing the areas of focus for their policy review. The systems and processes identified in this review can also assist policy-makers in other countries to safely implement health screening programmes within their contexts.

References

1. Holland WW, Stewart S, Masseria C. Policy Brief: Screening in Europe. European Observatory on Health Systems and Policies: World Health Organization; 2006.
2. Spiby J. International Comparison of Screening Process, Policy and Programmes. Unpublished.
3. Wilson JMG, Jungner G. Principles and practice of screening for disease. Geneva: World Health Organisation; 1968.
4. Policy review process and criteria for appraising the viability, effectiveness and appropriateness of a screening programme. [cited January 2014]; Available from: http://www.screening.nhs.uk/policyreview

http://www.screening.nhs.uk/criteria
5. Holland WW, Stewart S. Screening practices in Europe. In: Holland WW, Stewart S, editors. Screening for disease prevention: What works? Oxford, Seattle: Radcliffe Publishing Ltd; 2005. p. 139-42.
6. Andermann A, Blancquaert I, Beauchampb S, Déryc V. Revisiting Wilson and Jungner in the genomic age: a review of screening criteria over the past 40 years. Bulletin of the World Health Organization. 2008; **86**(4): 317-19.
7. Blancquaert I, Beauchamp S, Andermann A. Decision support guide for population-based genetic screening; 2009.
8. Andermann A, Blancquaert I, Beauchamp S, Costea I. Guiding Policy Decisions for Genetic Screening: Developing a Systematic and Transparent Approach. Public Health Genomics. 2009.
9. Andermann A, Blancquaert I, Dery V. Genetic screening: a conceptual framework for programmes and policy-making. Journal of Health Services Research & Policy 2010; **15** (2): 90–97.
10. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 2009; **6**(7): e1000097.
11. Rodriguez-Granger J, Alvargonzalez JC, Berardi A, Berner R, Kunze M, Hufnagel M, et al. Prevention of group B streptococcal neonatal disease revisited. the DEVANI European project. European Journal of Clinical Microbiology and Infectious Diseases. 2012; **31**(9): 2097-104.
12. Melin P, Verschraegen G, Mahieu L, Claeys G, De Mol P. Towards a Belgian consensus for prevention of perinatal group B streptococcal disease. Indian Journal of Medical Research, Supplement. 2004; **119**: 197-200.
13. Money DM, Dobson S. The prevention of early-onset neonatal group B streptococcal disease. Journal of obstetrics and gynaecology Canada: JOGC = Journal d'obstetrique et gynecologie du Canada: JOGC. 2004; **26**(9): 826-40.
14. Berardi A, Lugli L, Baronciani D, Creti R, Rossi K, Ciccia M, et al. Group B streptococcal infections in a northern region of Italy. Pediatrics. 2007; **120**(3): e487-93.
15. Melin P. Prevention of perinatal group B streptococcal diseases: Belgian guidelines. Round Table Series - Royal Society of Medicine. 2007; (85): 29-41.
16. Verani JR, McGee L, Schrag SJ. Prevention of perinatal group B streptococcal disease--revised guidelines from CDC, 2010. MMWR Recomm Rep. 2010; **59**(RR-10): 1-36.
17. Charter of the Czech Gynaecological and Obstetrical Society. [cited May 2014]; Available from: http://www.cgps.cz/cgps/stanovy.php
18. Campbell N, Eddy A, Darlow B, Stone P, Grimwood K. The prevention of early-onset neonatal group B streptococcus infection: technical report from the New Zealand GBS Consensus Working Party. The New Zealand Medical Journal **117**(1200).

19. Antenatal prevention of early neonatal bacterial infection - Clinical Practice Guidelines. France: Agence Nationale d'Accre´ditation et d'Evaluation en Sante, Guidelines Department 2001.
20. Harris R, Helfand M, Woolf S, Lohr K, Mulrow C, Teutsch S, et al. Current methods of the U.S. Preventive Services Task Force: a review of the process. Am J Prev Med. 2001; **20**(3S): 21–35.
21. Shah V, Ohlsson A, CTFPHC. Prevention of early-onset group B streptococcal infection in the newborn: Systematic review and recommendations. London, Ontario, Canada: Canadian Task Force on Preventive Health Care Technical Report 2001.
22. Prevention of perinatal group B streptococcal infections - Guidelines. Belgium: Belgian Health Council; 2003.
23. Policy directive: neonatal minimisation of early onset group B streptococcal (EOGBS) infection: Department of Health, NSW; 2005.
24. Screening and Treatment for Group B Streptococcus in Pregnancy, RANZCOG statement C-Obs 19. Melbourne: Royal Australian and New Zealand College of Obstetricians and Gynaecologists; 2007.
25. Prevention of Perinatal Group B Streptococcal Disease: Dutch Organisation of Obstetrics and Gynaecology (NVOG); 2008.
26. South Australian Perinatal Practice Guidelines: prevention and treatment of neonatal sepsis (including maternal group B streptococcal colonisation). Australia: Department of Health, Government of South Australia; 2010.
27. Supplement: Early onset Group B streptococcal disease. Australia: Queensland Government; 2011.
28. Hisanori M, Hiramatsu Y, Mitsuhiko K, Tomoyuki F, Hiromi H, Yoshinori I, et al. Guidelines for obstetrical practce in Japan: Japan Society of Obstetrics and Gynecology; 2011.
29. SEGO, SEN, SEIMC, SEQ, SEMFC. Prevention of perinatal group B streptococcal infections: revised Spanish recommendations 2012. Prog Obstet Ginecol. 2012; **55**(7): 337—46.
30. Evidence-Based Recommendations—GRADE. [cited January 2014]; Available from: http://www.cdc.gov/vaccines/acip/recs/GRADE/about-grade.html - resources
31. Betreuung der Neugeborenen von Müttern, die mit Streptokokken der Gruppe B kolonisiert sind: Swiss society of neonatology.
32. Prevention: Population Screening Act. [cited January 2014]; Available from: http://www.gezondheidsraad.nl/en/publications/prevention/results/taxonomy%3A239
33. Population surveys and screening. [cited January 2014]; Available from: http://www.rivm.nl/Onderwerpen/B/Bevolkingsonderzoeken_en_screeningen/Achtergrondinformatie
34. Screening to Improve Health in New Zealand: Criteria to assess screening programmes. New Zealand: National Advisory Committee on Health and Disability; 2003.
35. How to judge a proposal for a screening programme. France: Agence nationale d'accréditation et d'évaluation en santé; 2004.
36. Raccomandazioni per la pianificazione e l'esecuzione degli screening di popolazione per la prevenzione del cancro della mammella, del cancro della cervice uterina e del cancro del colon retto. Italy: The National Observatory of Screening; 2005.
37. Strong K, Wald N, Miller A, Alwan A. Current concepts in screening for noncommunicable disease: World Health Organization Consultation Group Report on methodology of noncommunicable disease screening. J Med Screen. 2005; **12**: 12–19.
38. Government Decree on Screenings. Finland: Ministry of Social Affairs and Health; 2006.

39. U.S. PreventiveTask Force Procedural Manual United States of America: U.S. Preventive Task Force; 2008.
40. Prophylaxe der Neugeborensepsis - frühe Form - durch Streptokokken der Gruppe B. Germany: Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften; 2008.
41. Early detection / prevention. 2009 [cited January 2014]; Available from: http://www.g-ba.de/institution/themenschwerpunkte/frueherkennung/ https://http://www.g-ba.de/institution/themenschwerpunkte/methodenbewertung/frueherkennung/
42. Documento marco sobre cribado poblacional: Ponencia de Cribado Poblacional de la Comision de Salud Publica; 2010.
43. Zwahlen M, Low N, Borisch B, Egger M, Obrist R, Paccaud F, et al. Swiss Society for Public Health: Workshop « A national screening Committee for Switzerland? Swiss Medical weekly (submitted). 2010.
44. Canadian Task Force on Preventive Health Care Procedure Manual. Canada: Canadian Task Force on Preventive Health Care; 2011.
45. Grades of Recommendation, Assessment, Development, and Evaluation: Canadian Task Force on Preventive Health Care; 2011.
46. USPSTF A and B Recommendations. . [cited; Available from: http://www.uspreventiveservicestaskforce.org/uspstf/uspsabrecs.htm
47. Holland WW. Screening for disease – considerations for policy. Euro Observer. 2006; **8**(3): 1-4.
48. Screening: a report 2001 [cited January 2014]; Available from: http://etiskraad.dk/upload/publications-en/screening/screening/ren.htm
49. Banta D, Oortwijn W, Cranovsky R. Health Policy, health technology assessment, and screening in Europe. International Journal of Technology Assessment in Health Care. 2001; **17**(3): 409–17.
50. Council Recommendation of 2 December 2003 on cancer screening (2003/878/EC). Official Journal of the European Union. 2003; **L327/34**.
51. Kristensen FB, Sigmund H. Health Technology Assessment Handbook Danish Centre for Health Technology Assessment, National Board of Health; 2007.
52. Population Based Screening Framework: Australian Population Health Development Principal Committee, Screening Subcommittee; 2008.
53. National Screening Unit Strategic Plan 2010 to 2015: National Screening Unit; 2010.
54. Harris R, Sawaya GF, Moyer VA, Calonge N. Reconsidering the Criteria for Evaluating Proposed Screening Programs: Reflections From 4 Current and Former Members of the U.S. Preventive Services Task Force. Epidemiologic Reviews. 2011; **33**: 20-35.
55. Nationella screeningprogram– modell för bedömning, införande och uppföljning. Sweden: National Board of Health and Welfare 2013.
56. Anttila A, Martin-Moreno JM. Cancer screening in Successes and Failures of Health Policy in Europe. In: Mackenbach JP, McKee M, editors. Successes and Failures of Health Policy in Europe: Four decades of divergent trends and converging challenges. New York: Open University Press; 2013. p. 179-92.
57. Streetly A, Elhers L. Population screening and public health. In press.
58. Schünemann H, Brozek J, Oxman A, editors. GRADE handbook for grading the quality of evidence and the strength of recommendation: GRADE Working Group; 2009.
59. Genetic screening programmes: an international review of assessment criteria: PHG Foundation; 2014.
60. Genetic Screening. The Hague: Health Council of the Netherlands: Committee Genetic Screening, Health Council; 1994.

61. On behalf of the European Society of Human Genetics' PPPC. Population genetic screening programmes: technical, social and ethical issues. European Journal of Human Genetics. 2003; **11**: 903-05.
62. 25 Recommendations on ethical, legal, and social implications of genetic testing. Brussels: European Commission; 2004.
63. Bonham JR. Impact of new screening technologies: should we screen and does phenotype influence this decision? Journal of inherited metabolic disease. 2013; **36**: 681-6.
64. Burton H, Moorthie S. Expanded newborn screening, a review of the evidence: PHG Foundation; 2010.
65. Goel V. Appraising organised screening programmes for testing for genetic susceptibility to cancer. BMJ. 2001; **322**(7295): 1174-78.
66. Sankaranarayanan R, Budukh AM, Rajkumar R. Effective screening programmes for cervical cancer in low- and middle-income developing countries. Bulletin of the World Health Organization. 2001; **79**(10): 954-62.

Appendices

Search strategies

General/Group B streptococcus screening

Medline

1	exp Mass Screening/
2	exp Policy Making/ or Public Policy/ or Health Policy/
3	exp Guideline/
4	exp Decision Making/
5	exp "review"/
6	health planning/ or exp health planning guidelines/ or exp health planning technical assistance/ or regional health planning/
7	(decision making* or decision-making*).kw,ti.
8	National Health Programs/ or Government Programs/
9	"screen*".kw,ti.
10	1 or 9
11	exp Streptococcus agalactiae/
12	"group b streptococc*".kw,ti.
13	"streptococc* agalactiae".kw,ti.
14	11 or 12 or 13
15	(polic* or guideline* or program* or strateg* or decision making* or decision-making* or process* or procedure* or review* or plan* or recommend* or committee*).ab,kw,ti.
16	exp Government Agencies/
17	2 or 3 or 4 or 5 or 6 or 7 or 8 or 15 or 16
18	10 and 14 and 17
19	limit 18 to yr="1996 -Current"

Embase

1	exp screening/
2	policy/
3	exp health care policy/
4	exp hospital policy/
5	exp practice guideline/
6	exp health program/
7	decision making/
8	process design/ or process development/ or process optimization/
9	procedures/
10	"review"/
11	hospital planning/ or patient care planning/ or planning/ or strategic planning/ or health care planning/
12	program development/
13	exp advisory committee/
14	"screen*".ti,kw.
15	1 or 14
16	(polic* or guideline* or program* or strateg* or decision making* or decision-making* or process* or procedure* or review* or plan* or recommend* or committee*).ti,ab,kw.
17	exp consensus development/
18	2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 16 or 17
19	exp Streptococcus agalactiae/
20	"streptococc* agalactiae".ti,ab,kw.
21	"group b streptococc*".ti,ab,kw.
22	19 or 20 or 21
23	15 and 18 and 22
24	limit 23 to yr="1996 -Current"

ASSIA

screen* AND (streptococci* agalactiae OR group b streptococci*) AND (polic* OR guideline* OR strategy* OR program* OR decision making OR decisionmaking OR process* OR procedure* OR review* OR plan* OR recommend* OR committee*)

1. 1996-2013

SSCI

Title=(screen*) AND Title=(group b streptococc* OR streptococc* agalactiae) AND Topic=(polic* or guideline* or program* or strateg* or decision making* or decision-making* or process* or procedure* or review* or plan* or recommend* or committee*)

Timespan=1996-2013. Databases=SCI-EXPANDED.

Genetic screening

PubMed	**Embase**
1. Mass screening 2. Policy making or public policy or health policy 3. Guideline 4. Decision making 5. review 6. Health planning or health planning guidelines or health planning technical assistance or regional health planning 7. Decision making or decision-making*Ti/Abs 8. national health programs or government programs 9. Screen* Ti/Abs 10. 1 or 9 11. Genetic* Ti/Abs 12. Genomic* Ti/Abs 13. 11 or 12 14. polic*OR guideline*OR program*OR strateg* OR process* OR procedure*OR review*OR plan*OR recommend*OR committee* Ti/Abs 15. government agencies 16. 2 or 3 or 4 or 5or 6or 7or 8 or 14 or 15 17. 10 and 13 and 16 18. Limit 18 to yr#1996-current	1. screening 2. policy 3. health care policy 4. hospital policy 5. practice guideline 6. health program 7. decision making 8. process design or process development or process optimization 9. procedures 10. "review" 11. hospital planning or patient care planning or planning or strategic planning or health care planning 12. program development 13. advisory committee 14. "screen*".ti,kw. 15. 1 or 14 16. (polic* or guideline* or program* or strateg* or decision making* or decision-making* or process* or procedure* or review* or plan* or recommend* or committee*).ti,ab,kw. 17. consensus development 18. 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 16 or 17 19. "genetic*".ti,ab,kw 20. "genomic*".ti,ab,kw. 21. 19 or 20 22. 15 and 18 and 21 23. limit 23 to yr="1996 -Current"
Social Science Citation Index (SSCI) Title=(screen*) AND Title=(genetic* OR genomic*) AND Topic=(polic* or guideline* or program* or strateg* or decision making* or decision-making* or process* or procedure* or review* or plan* or recommend* or committee*) Timespan=1996-2013. Databases=SCI-EXPANDED.	**Applied Social Science Index and Abstracts (ASSIA)** screen* AND (genetic* OR genomic*) AND (polic* OR guideline* OR strategy* OR program* OR decision making OR decisionmaking OR process* OR procedure* OR review* OR plan* OR recommend* OR committee*)

Examples of supporting evidence guides

Blancquaert et al. (2009)[7] Supporting evidence for criteria in decision node A

Supporting evidence

- Achievement of a consensus on the target objectives.
- Detailed description of each alternate strategy under consideration (i.e. which condition will be screened for, which test or series of screening tests will be used, which interventions will be offered and in which individuals will screening be indicated), including the status quo.

- Availability of epidemiological data for the population and for high-risk subgroups receiving standard clinical care establishing:
  - the incidence and prevalence
  - the average or median age of symptom onset
  - the mortality, morbidity, quality of life and disability
  - the utilization rate of health care services
  - the impact of the disease on the capacity for social adjustment or integration.

- As for criterion A2, plus:
  - the prevalence and attributable fraction of the risk factors being screened for
  - the cumulative risk of developing the disease and its complications for different combinations of risk factors.

- Identification of a window of opportunity for intervening after detection of the latent or early phase or of a risk factor, taking into consideration the timing of screening, disease progression, and ability to intervene.

- Availability of epidemiological data for the population and for high-risk subgroups undergoing screening establishing:
  - the analytical validity of screening tests for the disease or the risk factors to be detected (precision, detection threshold, reproducibility, sensitivity, specificity, PPV, NPV)
  - the clinical validity applicable to the target population (sensitivity, specificity, PPV, NPV).
- Choice of threshold values and compromises between sensitivity and specificity justified on the basis of the distribution of results in the target population and the consequences of false-negative and false-positive results.
- Documentation of ease and practicality of use on a large scale.
- Likelihood and nature of harms linked to the test and the information it provides are known and considered to be acceptable.

- Relative efficacy, effectiveness and safety of the therapeutic or preventive measures, including those involving behaviour changes, in relation to standard clinical care and possible alternatives:
  - Reduced mortality (e.g., disease-specific mortality rates, life expectancy, potential years of life lost)
  - Reduced morbidity or disability
  - Improved quality of life or capacity for social adjustment or integration
  - Improved summary measures (e.g. disability-adjusted life expectancy, QALY)
  - Likelihood and nature of harms linked to the intervention both known and acceptable.
- Reproductive options for the current pregnancy, future pregnancies, and for relatives:
  - Greater reproductive confidence within high-risk families
  - Improved quality of life or capacity for social adjustment or integration for families
  - Impact on family dynamics
  - Satisfaction of individuals and their families

- Have the most relevant alternative diagnostic or screening strategies for the condition been considered, in terms of target population, tests and interventions?
- Has the full range of potential benefits and risks to individuals and families been considered?
  - Potential benefits: reduced mortality, morbidity or disability; improved quality of life or capacity for social adjustment or integration; greater reproductive confidence.
  - Potential risks: consequences of false positive and false negative results (e.g., unnecessary procedures, inappropriate reassurance); anxiety; incidental findings (including non-paternity status and supplementary risk information that was not intended as an objective of screening); impact on family dynamics; disclosure of information to third parties; stigmatization, discrimination, or exclusion.
- Are the benefits and risks hypothetical or well-documented?
- Have the individuals or groups who will benefit from screening and those who will be exposed to the risks been clearly identified?
- Have any conflicts between criteria and principles been noted?
- Have the various stakeholders been identified and are their positions known?
- Has the appraisal of each criterion been documented?
- Have the trade-offs between different considerations (e.g., benefits to individuals vs. families) been clarified?

Harris et al. (2011)[54] Minimal evidence for calculating benefits and harms

Table 4. Minimal Evidence Sufficient to Estimate Benefits and Harms of Screening Programs With at Least Moderate Certainty

Issue	Minimal Evidence Required	Comments
	Benefits	
Probability of an adverse health outcome in the population without screening	1. Minimally selected population from a broad setting, reasonably representative of the population to be screened. 2. Cohort study with a reasonable measure of the health outcome over time. 3. Well-conducted retrospective studies may be useful.	1. It is important to have a reasonable estimate of the *absolute magnitude* of the incidence and severity of the adverse health outcome, because they will be needed for the outcomes table. 2. The more highly selected the population, the less useful the study. 3. Ideally, needed are several cohort studies from different populations with similar results.
The degree to which screening identifies all people who would suffer the adverse health outcome	1. Incidence of interval cases and stage distribution of cases found and not found by screening within a properly controlled trial. 2. Longitudinal observational studies demonstrating the ability of the testing strategy to correctly categorize the population into those who will suffer the adverse health outcome and those who will not, over an important period of time. 3. Retrospective studies potentially useful if well conducted. 4. Studies should be generalizable to the general population.	1. Cross-sectional studies are not useful. 2. Important to have a reasonable estimate of the *absolute magnitude* of incidence or progression. 3. Ideally, needed are several studies from different populations with similar results.
Magnitude of the incremental health benefit of earlier versus later treatment resulting from screening	1. Well-conducted randomized controlled trials (or a well-conducted nonrandomized controlled trial) of treatment if participants (or a defined and separately analyzed subset) are found by screening, previously undiagnosed and asymptomatic. 2. Populations studied should be reasonably representative of the population to be screened. 3. Dramatic effect (or no effect) in concurrent nonrandomized controlled studies with similar populations, some screened and some unscreened, may be useful if well conducted. 4. Adequately controlled trials with intermediate outcome if the intermediate outcome is closely associated with the health outcome, especially if evidence is consistent from several studies.	1. Ideal evidence is a randomized controlled trial of screening, although adequately controlled (not historical controls) nonrandomized trials may be sufficient if well conducted and in representative populations. 2. Case-control, cohort, ecologic, and modeling evidence alone is insufficient. 3. Evidence is stronger if consistent evidence is obtained from different types of studies in different populations. 4. Cohort and modeling studies may be adequate to extend results from controlled trials to other populations or to help determine frequency of screening or starting/stopping ages for screening.

Table continues

USA strength of recommendation

USPSTF Strength of evidence				
Certainty of Net Benefit	**Magnitude of Net Benefit**			
Substantial	**Moderate**	**Small**	**Zero/negative**	
High	**A**	**B**	**C**	**D**
Moderate	**B**	**B**	**C**	**D**
Low	**Insufficient**			

Grade	Grade Definitions	Suggestions for Practice
A	The USPSTF recommends the service. There is high certainty that the net benefit is substantial.	Offer/provide this service.
B	The USPSTF recommends the service. There is high certainty that the net benefit is moderate or there is moderate certainty that the net benefit is moderate to substantial.	Offer/provide this service.
C	The USPSTF recommends against routinely providing the service. There may be considerations that support providing the service in an individual patient. There is moderate or high certainty that the net benefit is small.	Offer/provide this service only if there are other considerations in support of the offering/providing the service in an individual patient.
D	The USPSTF recommends against the service. There is moderate or high certainty that the service has no net benefit or that the harms outweigh the benefits.	Discourage the use of this service.
I Statement	The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of the service. Evidence is lacking, of poor quality or conflicting, and the balance of benefits and harms cannot be determined.	Read "Clinical Considerations" section of USPSTF Recommendation Statement. If offered, patients should understand the uncertainty about the balance of benefits and harms.