Newborn screening for biotinidase deficiency

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

Version: Final

Author: Bazian

Date: July 2017

The UK National Screening Committee secretariat is hosted by Public Health England.

About the UK National Screening Committee (UK NSC)

The UK NSC advises ministers and the NHS in the 4 UK countries about all aspects of population screening and supports implementation of screening programmes.
Conditions are reviewed against evidence review criteria according to the UK NSC's evidence review process.

Read a complete list of UK NSC recommendations.

UK NSC, Floor 2, Zone B, Skipton House, 80 London Road, London SE1 6LH
www.gov.uk/uknsc
Twitter: @PHE_Screening Blog: phescreening.blog.gov.uk

For queries relating to this document, please contact:
phe.screeninghelpdesk@nhs.net

Published February 2018

Contents

Plain English summary

This review aimed to look at whether screening for biotinidase deficiency should be offered as part of the newborn blood spot screening programme.

Biotinidase is an enzyme that the body needs to recycle a vitamin called biotin. Biotinidase deficiency is a rare genetic condition where people do not make enough of this enzyme. Babies can inherit the condition from their parents if both parents carry a mutation in the biotinidase gene. There are two forms of the condition: a severe or “profound” deficiency and a milder “partial” deficiency.

People with severe (profound) deficiency usually have biotinidase that works at less than 10% of its normal level. Without treatment they can develop problems with their nervous system, including vision and hearing loss, coordination problems and seizures. People with the milder (partial) deficiency have biotinidase that works at 10% to 30% of the normal level. They may only develop symptoms if they have other illnesses, like an infection.

Treatment involves taking biotin by mouth throughout life. Evidence from the previous UK National Screening Committee (NSC) report suggested that biotin treatment is highly effective with no known side effects.

The UK NSC does not currently recommend newborn screening for biotinidase deficiency. The screening test itself is fairly simple in that it involves testing enzyme activity in a dried blood spot sample, which is already collected to screen for other conditions. However, the last UK NSC review in 2012 found several gaps in the evidence. There was no information on how common biotinidase deficiency was in the UK. The level of biortinidase activity that should be used to diagnose biotinidase deficiency was unclear. Everyone diagnosed with biotinidase deficiency usually starts biotin treatment at the time of diagnosis. However, some people with partial deficiency do not develop symptoms so it is questionable whether everyone needs treatment.

The current review looked to see whether the evidence published since 2012 has answered these questions, but found that it had not.

There is still no data on how common biotinidase deficiency is in the UK. Screening programmes from European and North America have shown that the number of cases can vary widely from 1 in 4,508 to 1 in 31,717 people. Prevalence will vary by country according to ethnicity and likelihood of marriage between relatives. Therefore it is not possible to know the UK prevalence based on these figures.

Global screening programmes differ in the enzyme levels that they use to diagnose biotinidase deficiency, but roughly they show that only about half of newborns detected by the screening test have the condition. This means that because the test does not work well enough around half of babies who screen positive do not have the condition. This might result in unnecessary stress for the families of babies who have to have further testing. Furthermore, only around one in 10 babies with positive screening tests will have the more severe (profound) form of the deficiency. This might result in treating babies that do not need it. It is also unclear how many newborns with biotinidase deficiency may be missed by the screening test, as those who screen negative are not monitored to see if they develop the condition later on.

It is still uncertain whether all screen-detected children need biotin treatment, particularly those with milder partial deficiency. The best dose of biotin to give is also unclear. Other countries that screen for biotinidase deficiency have differed in the dose they prescribe for diagnosed children. Most children have remained symptom-free with treatment, but it is not known whether some would have remained symptom-free without biotin. There are reports of individuals who have developed symptoms despite taking biotin, but there is no clear pattern to who develops these symptoms. Some symptoms, such as skin rashes, may not have been caused by the disease. There are also a few reports of people who have stopped treatment, but again there is no consistent pattern in what happens to them. Some people with profound or partial deficiency have remained symptom-free without treatment while others have developed symptoms.

As these various uncertainties remain, the evidence does not suggest that the current policy not to perform newborn screening for biotinidase deficiency should be changed at the current time.

Executive summary

Purpose of the review

This review aimed to see whether there is evidence to support newborn screening for biotinidase deficiency.

Background

Biotinidase deficiency is a rare autosomal recessive genetic condition where there is reduced activity of the enzyme that recycles biotin.

There are two levels of biotinidase deficiency: profound (severe) deficiency and partial (milder) deficiency. Around half of those affected have profound (severe) deficiency, defined as less than 10% of normal biotinidase activity. They may present with symptoms such as seizures, ataxia, visual and hearing difficulties. Partial biotinidase deficiency is a milder form usually defined by having 10% to 30% of normal biotinidase activity. People with partial deficiency may only present with symptoms during periods of metabolic stress. The clinical course varies widely within and between the two types.

Treatment for both types is oral supplementation with unbound (free) biotin. The last UK National Screening Committee (NSC) review reported that biotin is regarded as a highly effective treatment with no known adverse effects. If relying on symptomatic presentation, there may be substantial delay (years) in diagnosis which could lead to poorer outcomes.

However, there are various uncertainties around treatment, such as the appropriate biotin dose for partial biotinidase deficiency.

Recommendation under review

The UK NSC does not currently recommend newborn screening for biotinidase deficiency. The last external evidence review conducted in 2012 identified several uncertainties.

The screening test involves measuring biotinidase activity in a newborn dried blood spot (DBS) sample, so is relatively simple to perform given that newborn blood spots are already collected as part of the newborn screening programme. However, countries have differed in the enzyme activity cut-off used, and there has been limited test performance data.

The last review found no data on UK prevalence. There was also limited understanding of which screen-detected children with profound or partial biotinidase deficiency would go on to develop symptoms and therefore need treatment. Oral biotin is considered to be a safe and effective treatment and children with profound deficiency are always treated. However, there was more uncertainty around the management of partial deficiency. Most children are treated, but there the dose given has varied between global treatment centres.

Focus of the review

This review aims to see whether evidence has been published since 2012 that addresses these uncertainties and suggests a need to reconsider the screening policy.

The review addresses 4 key questions:

- the prevalence of biotinidase deficiency in the UK population
- the natural history of profound and partial biotinidase deficiency
- whether a screening test cut-off has been identified
- the treatment outcomes of people with profound and partial deficiency

Findings and gaps in the evidence of this review

The review did not identify the evidence needed to answer these questions:

1 There is still no UK incidence/prevalence data available. Screening programmes and national screening pilots from North America, Europe and the United Arab Emirates (UAE) find highly variable incidence, ranging from 1 in 4,508 to 1 in 31,717 for biotinidase deficiency overall, and between 1 in 25,000 and 1 in 100,000 for profound deficiency. Prevalence is likely to be influenced by ethnic variation and levels of consanguinity. Therefore it is difficult to extrapolate this data to the UK.
2 The majority of children diagnosed with partial or profound deficiency are treated with biotin. Therefore, there is no data to inform the clinical course of untreated profound or partial deficiency (by enzyme activity or genotype) and explain why some people remain asymptomatic. Consequently, it is still not clear if all screen-detected children need treatment.
3 Uncertainties remain around the optimal enzyme activity threshold to use in newborn DBS screening. North American and European screening programmes show that DBS enzyme activity has a fairly low but variable PPV for biotinidase deficiency, ranging from 30% to 86%. Most studies suggest that fewer than 1 in 10 screen positives will have profound deficiency, which could have implications given the uncertainties over treatment of partial deficiency. The variation in PPV across programmes is likely explained by differences in population prevalence and characteristics of the screening test. Enzyme activity increases during the first days to weeks of life. The cohort with the best PPV (86%), despite lowest population prevalence, performed a repeat DBS for those in the partial range of enzyme activity. Other cohorts have decreased their DBS enzyme activity threshold in recent years to improve screening test accuracy (such as changing from <30% to <20%). The optimal screening test threshold and/or timing (such as performing a later repeat DBS for those with partial levels) remain to be clarified. There is no follow-up of screen negatives, so no further test performance data (such as sensitivity and specificity) available.
4 As most children are treated at diagnosis, RCTs or comparative studies comparing treated and untreated populations are not available. Cohorts of children from North American and European screening programmes suggest that most people remain asymptomatic on biotin. The biotin dose prescribed has been variable. However, it is not known who would have remained asymptomatic without treatment. Reports of symptoms occurring while on treatment

have been inconsistent, both for partial and profound deficiency. It is unclear whether all are disease-related. Similarly reported effects of treatment non-compliance both for partial and profound deficiency are inconsistent, with some people developing symptoms and others remaining asymptomatic. The evidence is not available to inform which screen-detected children with partial or profound deficiency would develop symptoms and need biotin treatment, or the optimal dose to give. Neither can the evidence inform whether screen detection improves outcomes compared with clinical detection.

Recommendations on screening

The evidence needed to support a screening programme for biotinidase deficiency in the UK is not available. The findings indicate that a change to the current policy should not be made and screening for biotinidase deficiency should not be recommended.

It may be difficult to obtain the evidence needed to meet the criteria for a screening programme. Children are treated at diagnosis, so study of natural history without treatment would be difficult. Also obtaining further test performance data (such as the rate of false negatives) would be difficult given that biotinidase deficiency is rare. A non-selected population screening cohort that is large enough to detect sufficient cases, while performing confirmatory re-testing for all screen negatives, is unlikely.

Nevertheless, further research would be valuable to:

- establish the incidence/prevalence of biotinidase deficiency in the UK
- identify which levels of enzyme activity or genotypes confer increased likelihood of adverse outcomes, and which may remain asymptomatic
- better clarify the optimal threshold to use in a DBS screening test, and determine whether lower enzyme activity thresholds, or a repeat screen, could preferentially identify those who are most likely to develop symptoms and would benefit from earlier screen detection and treatment
- clarify the biotin dose to use for partial or profound deficiency

Limitations

This was a rapid evidence review process. Searching was limited to 3 bibliographic databases and did not include grey literature sources.

Literature search and initial appraisal were undertaken by one information specialist. Second pass appraisal and study selection was then conducted by 2 analysts. Any decisions on study inclusions, queries or scope refinement were resolved in a meeting between analysts and in discussion with UK NSC evidence team as needed. Systematic reviews, RCTs and prospective controlled studies were prioritised (if available) before moving down through the lower hierarchy of evidence.

We did not include studies that were not available in English language, and did not review conference abstracts, conference reports or poster presentations. We were also unable to contact study authors or review non-published material.

Introduction and approach

Background

Biotinidase is an enzyme that recycles biotin that is encoded by the *Biotinidase* (*BTD*) gene on chromosome 3q25. Biotinidase deficiency is a rare autosomal recessive genetic condition where there is reduced activity of this enzyme.

Around half of those with biotinidase deficiency have profound (severe) deficiency, normally defined as less than 10% of normal biotinidase enzyme activity.[1] About 150 different mutations on the *BTD* gene have been identified in individuals with profound deficiency, though 5 common mutations account for the majority (60%) of cases.[1]

Partial biotinidase deficiency is a milder form usually defined by 10% to 30% of normal biotinidase activity. Almost all (98%) people with partial biotinidase deficiency carry the 1330G>C (D444H) mutation as one of their alleles, along with a profound deficiency allele as the other.[1]

Children with profound deficiency usually present with symptoms, such as seizures, hypotonia, skin rash or alopecia, by around 2 to 5 months of age. Without treatment most will go on to develop hearing and visual loss, ataxia and developmental delay. People with partial deficiency may present with symptoms, such as hypotonia, skin rashes and hair loss, only during periods of metabolic stress, for example illness or infection.[1]

Treatment is oral supplementation with unbound (free) biotin. The last UK NSC review reported this to be a highly effective treatment with no known adverse effects. There may be substantial delay in diagnosis if relying on symptomatic presentations, and observational study has suggested this delay leads to poorer outcomes.[1] Screen-detection could prevent symptoms in pre-symptomatic children and lead to clinical improvement in those displaying early symptoms.

The test for biotinidase deficiency is analysing serum biotinidase activity level, which can be assessed simply and safely in the newborn DBS test.

However, the last review noted the absence of evidence-based policies on who should be treated.[1] All children with profound deficiency are treated with oral biotin at diagnosis, as are most with partial deficiency. However, the clinical course of the condition has been observed to vary widely both within and between the two types. As said, many with partial deficiency may only develop symptoms during times of metabolic stress. There have also been reported cases of people with profound deficiency who have remained asymptomatic into adulthood despite having no treatment.[1] Treatment centres also vary in the biotin doses used. The last review identified doses between 5 and 30mg biotin daily being prescribed for profound biotinidase deficiency, and between 1 and 10mg for patients with partial biotinidase deficiency.[1] Therefore many uncertainties around treatment and treatment need remain.

Current policy context and previous reviews

The UK NSC does not currently recommend newborn screening for biotinidase deficiency. This policy was last reviewed in 2012. An external evidence review[1] assessed the evidence published between 2004 and 2012 and identified several gaps in the literature:

1 The condition is known to be rare, but there was a lack of UK prevalence data.
2 There was limited understanding of how type of deficiency (partial or profound) or genotype is linked with symptomatic outcomes
3 The screening test is relatively simple and involves newborn DBS sampling. However, countries varied in what they considered to be a normal reference range for enzyme activity. Furthermore the sensitivity and number of false negatives was uncertain as follow-up testing was not performed for screen negatives.
4 A highly effective and safe treatment is available and all children with profound deficiency are treated. However, there was uncertainty around the management of partial deficiency; most children appear to be treated but there was variability in the dose given.

Objectives

The current review considers whether new evidence has been published since January 2012 to suggest that the current policy not to screen newborns for biotinidase deficiency should be reconsidered.

Four questions will be addressed to cover the uncertainties raised by the last external evidence review. These are presented in Table 1.

Table 1. Questions for the evidence summary, and relationship to UK NSC screening criteria

	Criterion	Questions	Studies Included
	THE CONDITION		
1	The condition should be an important health problem as judged by its frequency and/or severity. The epidemiology, incidence, prevalence and natural history of the condition should be understood, including development from latent to declared disease and/or there should be robust evidence about the association between the risk or disease marker and serious or treatable disease.	Q1: What is the UK prevalence or incidence of biotinidase deficiency?	6 cohorts
		Q2: What is the natural history of profound and partial deficiency?	0 studies
	THE TEST		
4	There should be a simple, safe, precise and validated screening test.	Q3: Has a cut-off been identified for biotinidase deficiency screening in newborns?	4 cohorts
	THE INTERVENTION		
9	There should be an effective intervention for patients identified through screening, with evidence that intervention at a pre-symptomatic phase leads to better outcomes for the screened individual compared with usual care. Evidence relating to wider benefits of screening, for example those relating to family members, should be taken into account where available. However, where there is no prospect of benefit for the individual screened then the screening	Q4: What treatment outcomes have been reported for those with partial and profound deficiency?	4 cohorts

Criterion	Questions	Studies Included
programme should not be further considered.		

Methods

The current review was conducted by Bazian, in keeping with the UK NSC evidence review process. Database searches were conducted on 12 and 13 April 2017 to identify studies relevant to the questions detailed in Table 1.

Eligibility for inclusion in the review

A systematic literature search of three databases was performed for studies published between January 2011 and April 2017.

After de-duplication the search yielded 180 references addressing biotinidase deficiency. These studies were further filtered at title and abstract level by one information specialist, and 72 studies were considered relevant to newborn screening for biotinidase deficiency.

These abstracts were further reviewed against the inclusion criteria by a second reviewer. Due to the low volume of evidence identified, no restrictions were placed on study design or country of origin. Studies of potential relevance, or those where applicability was uncertain from the abstract, were selected for full text appraisal to ensure that all potentially relevant studies were captured.

Eighteen studies were selected for full text appraisal. Each full text article was reviewed against the inclusion/exclusion criteria by one main reviewer, who determined whether the article was relevant to one or more of the review questions. Any uncertainties around inclusion/exclusion were resolved through input from a second reviewer, followed by further discussion with UK NSC evidence team as needed.

Eligibility criteria for each question are presented in **Error! Reference source not found.** below. Any refinements to these criteria (e.g. need to move down the hierarchy of evidence), and further information on the evidence selection process for each key question, is discussed in the evidence description for each criterion in the report below.

Table 2. Inclusion and exclusion criteria for the key questions.

Question	Inclusion criteria							Exclusion criteria
	Population	**Target condition**	**Intervention**	**Reference Standard**	**Comparator**	**Outcome**	**Study type**	
1	The newborn or general UK population, or otherwise analogous Western population	Biotinidase deficiency (partial or profound)	NA	NA	NA	NA	Cross sectional, cohorts, registry data or SRs of these studies	Studies with sample size <500; conference abstracts; non-English language studies.
2	Newborns with partial or profound deficiency (screen or clinically detected). Ideally UK or analogous Western population	Biotinidase deficiency (partial or profound)	NA	NA	Children with other severity (partial vs. profound) Healthy children without biotinidase deficiency	Seizures Ataxia Vision/hearing difficulties Breathing difficulties Hypotonia Alopecia Skin rashes	Prospective comparative cohorts. Non-comparative cohorts/case series if above not available. SRs of these studies	Single case reports; conference abstracts; non-English language studies.

3	General newborn population. Ideally UK or other analogous Western population	Biotinidase deficiency (partial or profound)	Enzyme activity on newborn blood spot	Mutation analysis Parental testing Re-testing after biotin supplements	NA	Test accuracy, validity outcomes: sensitivity, specificity, PPV, NPV, likelihood ratios (+/-)	Diagnostic cohorts with performance data available (or where this can be calculated). SRs of these studies. Guidelines	Studies in children with known status; conference abstracts; non-English language studies; animal studies.
4	Children with partial or profound deficiency. Ideally screen-detected otherwise clinical-detected. Ideally UK or other analogous Western population	Biotinidase deficiency (partial or profound)	Biotin supplementation	NA	No supplement or alternative dose. Treated children with other severity (partial vs. profound)	Seizures Ataxia Vision/hearing difficulties Breathing difficulties Hypotonia Alopecia Skin rashes	RCTs or comparative cohorts. Non-comparative cohorts/case series if above not available. SRs of these studies.	Conference abstracts; non-English language studies; animal studies.

Appraisal for quality/risk of bias tool

Each criterion was summarised as ‘met’, ‘not met’ or ‘uncertain’ by considering the results of the included studies in light of the volume, quality, consistency and applicability of the body of evidence.

Several factors were assessed to determine the quality of the identified evidence, including study design and methodology, risk of bias, directness and applicability of the evidence. Factors that were determined to be pertinent to the quality of the body of evidence identified for each criterion are discussed in the results section for each question, and also outlined in the evidence extraction tables in Appendix 3.

Diagnostic accuracy studies considered for criterion 4 were assessed using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool. This quality assessment focused on 4 main domains: patient selection, the index test, the reference standard, and flow and timing of index test and reference standard. Each domain was assessed for risk of bias and applicability to a potential UK screening programme population. Results of these assessments are presented in Table 15 in Appendix 3.

Databases searched

Searches for the 4 questions were performed in MEDLINE and Embase databases (Embase.com) on 12 April, and The Cochrane Library (Wiley Online) on 13 April 2017. The full search strategy is presented in 0.

0contains a full PRISMA flow diagram (

Figure), along with a table of the included publications and details of which questions these publications were identified as being relevant to (Table 12. Summary o).

Question level synthesis

Criterion 1 – The condition should be an important health problem as judged by its frequency and/or severity. The epidemiology, incidence, prevalence and natural history of the condition should be understood, including development from latent to declared disease and/or there should be robust evidence about the association between the risk or disease marker and serious or treatable disease.

Question 1 – What is the UK prevalence or incidence of biotinidase deficiency?

Eligibility for inclusion in the review

The 2012 UK NSC external review[1] noted that the overall prevalence of biotinidase deficiency in Europe was estimated at 1 in 50,000. However, there was no data on incidence or prevalence in the UK.

This review aimed to identify cross sectional studies, cohorts or registry studies providing data on UK incidence or prevalence. If UK data was not available, analogous Western populations would be considered.

Description of the evidence

Database searches yielded 72 studies of relevance to the overall topic of biotinidase deficiency screening.

Forty-two of these studies were initially grouped as being relevant to the epidemiology (22) or natural history (20) questions. However, study groupings were only loosely assigned to question at initial appraisal, according to apparent title and abstract content.

At second appraisal, any studies of potential relevance to any of the 4 questions were selected for full text appraisal, regardless of study design or country of origin. Overall 18 of 72 studies were selected for full text appraisal. Nine of these studies came from the epidemiology and natural history groups. However, studies were often relevant to more than one question. Other studies initially considered relevant to the test and

treatment questions were ultimately found to contain incidence/ prevalence data at full text appraisal.

No UK incidence/prevalence data was identified.

Six studies contained incidence data from newborn screening programmes or pilots: one programme from The Netherlands,[2] one from Canada (Ontario),[3] two from US states[4, 5] and two national screening pilots from Greece[6] and the UAE.[7] Incidence data from these studies is summarised in Table 3, and a full data extraction is presented in the summary and appraisal of individual studies in **Error! Reference source not found.**. Most of these studies also provided data for other questions.

Appendix 2 contains a summary of the studies included and excluded at full text appraisal.

Summary of findings

Table 3. Incidence of biotinidase deficiency from cohorts of global screening programmes

Study	Setting	Population	Incidence
Wiltnik et al 2016[2]	Southwest Netherlands screening programme 2007 to 2012	Southwest n=304,982 screened n=50 confirmed (6 profound, 44 partial) Nationally n=913,965 screened n=111 confirmed (no distinction by type)	Southwest 1 in 6,100 1 in 50,830 profound* 1 in 6,931 partial* Nationally 1 in 8,233 (no distinction by type)
Gannavarapu et al 2015[3]	Ontario, Canada screening programme 2007 to 2014	n=1,043,895 screened n=71 confirmed (17 profound, 54 partial)	1 in 14,703* 1 in 61,406 profound* 1 in 19,331 partial*
Cowan et al 2012[4]	California, US screening programme 2007 to 2011	n=2,061,609 screened n=65 confirmed (n=28 profound, n=37 partial)	1 in 31,717 1 in 73,629 profound 1 in 55,719 partial*
Jay et al 2015[5]	Michigan, US screening programme 1988 to 2012	n screened not reported	1 in 14,539 1 in 101,779 profound 1 in 16,533 partial
Thodi et al 2013[6]	Greece screening pilot uncertain period	n=63,119 screened n=14 confirmed (all partial)	1 in 4,508 (overall and partial)
Al Hosani et al 2013[7]	UAE screening pilot 2010	n=25,000 screened n=3 confirmed (n=1 profound, n=2 partial)	1 in 8,300 1 in 25,000 profound* 1 in 12,500 partial*

* reviewer calculated

The 6 included studies provide incidence data based on confirmed cases identified through biotinidase deficiency screening programmes or pilot programmes planned for national enrolment.

Incidence was highly variable, ranging from 1 in 4,508 in Greece to 1 in 31,717 in California. This is much higher than the 1 in 50,000 estimate for Europe found in the last external evidence review.[1] The incidence of profound deficiency specifically is closer to that estimate, at between 1 in 50,000 and 1 in 100,000 in most studies (with the exception of the UAE cohort which had higher incidence at 1 in 25,000).

Sample size and representation could have contributed to the variable incidence. The Dutch,[2] US[4, 5] and Canadian[3] studies present data from established newborn screening programmes. They would be expected to give nationally or regionally representative figures (though participation rates aren't specifically discussed). It is possible that the smaller Greek[6] and UAE[7] pilots in particular may give less accurate estimates. These studies only include a proportion of the newborn population. The Greek[6] study reported only including newborns of Greek origin. The UAE[7] study reported including UAE nationals and represented around a third of all babies born in the UAE in 2010. No further eligibility criteria is given for these pilot studies so it is unclear how nationally representative they are. The smaller sample sizes could also make estimates less accurate and could explain the absence of profound cases in the Greek[6] study.

Cultural and ethnic variation may also explain a significant portion of the variation in prevalence globally. For example, Cowan et al[4] report that the majority of profound cases in their Californian sample (19/28) were in babies of Hispanic origin; Gannavarapu et al[3] report a highly ethnically diverse population being screened in Ontario; and the UAE may be expected to have a higher prevalence of consanguinity than many Western countries.

It is uncertain where UK incidence may lie in this spectrum.

The differences in diagnostic confirmation methods, for example laboratory analyser, reference range for enzyme activity (including ranges defined by enzyme response units rather than percentage biotinidase

activity), and whether this included parental enzyme testing or genotyping, may also have contributed to the variability in incidence.

Question 2 – What is the natural history of profound and partial biotinidase deficiency?

Eligibility for inclusion in the review

Although in general the natural history of biotinidase deficiency is reasonably well understood, there is limited understanding of how partial or profound deficiency (or associated genotype) correlates with clinical course. Disease progression can be variable. As discussed, individuals with partial deficiency usually remain asymptomatic but can develop symptoms during periods of metabolic stress. People with profound deficiency will usually develop neurological symptoms without treatment, but there have been reports of individuals with profound deficiency who have not developed symptoms. Therefore there is a need to better understand why some individuals remain asymptomatic.

This review aimed to identify comparative cohorts or case series that compared the clinical course between those with partial or profound deficiency, or those with particular genotype. Children in these studies would ideally be screen-detected and from populations analogous to the UK, but studies in clinically-detected individuals were also reviewed.

Description of the evidence

As for question 1 above, 18 studies in total were selected for full text appraisal, with no studies excluded at this stage based on country of origin or study design.

As almost all children diagnosed with biotinidase deficiency receive treatment, there was overlap between the studies relevant for this question – assessing the clinical course of partial and profound deficiency – and question 4 assessing treatment outcomes of partial and profound.

Four cohorts of relevance were identified. All included children identified through screening programmes who all received treatment. Two cohorts[2, 5] reported treatment response for children with both partial and profound deficiency, including some who discontinued treatment. Another cohort[8] described treatment response for children with profound deficiency only, some of whom discontinued treatment. One final screened cohort[3] described the effects of discontinuing treatment for three children with

partial deficiency. Though relevant to natural history, as all children were treated these studies were considered primarily relevant to criterion 9/question 4 and are discussed in that section.

One additional Polish study[9] just fell short of inclusion criteria. It assessed 22 children with biotinidase deficiency identified by varied means, most of whom were clinically detected, 2 were screen detected (Canadian and Turkish programmes), and a few were relatives subsequently identified. All received treatment, and the study describes response for those who were symptomatic or asymptomatic at diagnosis. Symptomatic or asymptomatic presentations did not correlate with level of enzyme activity or genetic mutation (that is, some with profound deficiency were asymptomatic at diagnosis while some with partial deficiency were symptomatic). Therefore the study was considered to have limited applicability to a screening programme as clinical course or treatment response could not be applied to a particular diagnostic type.

Individual case reports were excluded, as were studies reporting the mutation spectrum of patients but not looking at how this was associated with clinical course (that is, with no explicit assessment of genotype-phenotype correlation). Similarly studies reporting symptoms, but not describing the enzyme activity or mutation of these people and/or whether or not they were receiving treatment, were also excluded.

Summary of Findings Relevant to Criterion 1: not met.

Q1 Incidence/prevalence of biotinidase deficiency in the UK: no data was available. Incidence data from other countries or regions is highly variable ranging from 1 in 4,508 to 1 in 31,717, and between 1 in 25,000 and 1 in 101,779 for profound deficiency. This is higher than previous estimates for Europe. However, prevalence is likely to show ethnic and cultural variation, and therefore it is difficult to apply this data to infer what the UK incidence could be.

Q2 Natural history of profound and partial deficiency: no studies described the clinical course of profound or partial deficiency (by enzyme activity or genotype) in untreated populations. There is still no evidence to explain why some people remain asymptomatic, and treatment need following screen detection remains uncertain.

Criterion 4 – There should be a simple, safe, precise and validated screening test.

Question 3 – Has a cut-off been identified for biotinidase deficiency in newborns?

Eligibility for inclusion in the review

The 2012 UK NSC external evidence review[1] concluded how the screening test would be safe and simple as it uses the newborn blood spot already collected as part of the newborn screening programme. The screening test involves measuring enzyme activity using a DBS sample, which is then confirmed by a follow-up blood test. Profound deficiency is usually defined as biotinidase activity <10% normal and partial deficiency as 10% to 30% activity. However, countries in Europe and US states vary in the cut-offs used.

The last review[1] found a lack of performance data for the screening test. Not all screen positives received a confirmatory blood test, and a high false positive rate was reported. Neither was there follow-up of screen negatives. Therefore, sensitivity and specificity of the test could not be calculated.

This review aimed to see whether evidence assessing the performance of biotinidase screening tests, against confirmed diagnosis by mutation analysis or other reference standard in a non-selected (consecutive) sample of newborns, has been published since 2012.

Description of the evidence

Twenty-four studies were considered relevant to this question at initial appraisal, of which 6 were selected for full text appraisal. However, as described above, categories were only loosely assigned and other studies containing relevant data came from the total 18 studies reviewed at full text.

No studies had evaluated the performance of newborn screening tests with confirmatory re-testing of both screen positives and negatives. Therefore no data on sensitivity, specificity or negative predictive value (NPV) was available.

Neither have any studies assessed a screening test in a UK population.

However, 4 studies of screening programmes from Europe,[2, 6] Canada[3] and the US[4] gave the number of screen positives and confirmed diagnoses which allowed calculation of positive predictive value (PPV). These 4 studies are included for this question and are summarised in Table 4.

Studies assessing diagnostic tests (rather than blood spot screening) were excluded, as were analytical validity studies assessing screening tests in people with known diagnostic status. Other full text exclusions included studies where cut-off thresholds were not reported, and one study reporting how the recall rate in rural and inner city India decreased when the threshold was revised for the local population.

Appendix 2 contains a summary of the studies included and excluded at full text appraisal. A study-level summary of data extracted from each of these 4 publications is presented in the summary and appraisal of individual studies in **Error! Reference source not found.**.

Summary of findings

Table 4. PPV of biotinidase deficiency screening tests in global screening programmes

Study	Population and setting	Index test	Reference standard	Test performance data
Wiltnik et al[2] 2016	Netherlands screening programme, 2007 to 2012 Southwest n=304,982 Nationally n=913,965	Biotinidase activity <30% on DBS at mean 5 days (changed to <20% in July 2012)	Re-testing at mean 12 days (profound <10%, partial 10 to 30%) Genotype sequencing performed but not apparently required for confirmation	Southwest: n=92 screen positive n=50 confirmed (6 profound, 44 partial) PPV 54.3%* PPV for profound 6.5%* Nationally: n=261 screen positive n=111 confirmed PPV 42.5%*
Gannavarapu et al[3] 2015	Ontario, Canada screening programme	Biotinidase <10 ERU on DBS (changed to <35 MRU in 2014)	Re-testing (Collection time not given) Genotype	n=246 screen positive n=71 confirmed (17 profound, 54 partial)

Study	Population and setting	Index test	Reference standard	Test performance data
	2007 to 2014 n=1,043,895	(Collection time not given)	sequencing performed but not apparently required for confirmation	PPV 28.9%* PPV for profound 6.9%*
Cowan et al[4] 2012	California, US screening programme 2007 to 2011 n=2,061,609	Biotinidase ≤6 ERU on DBS, or 6 to 10 on two DBS (Collection time not given)	Re-testing with parental/unrelated relative testing (Collection time not given) Genotype sequencing performed but not apparently required for confirmation	n=76 screen positive n=65 confirmed (28 profound, 37 partial) PPV 85.5% PPV for profound 36.8%*
Thodi et al[6] 2013	Greece screening programme (uncertain period) n=63,119	Biotinidase activity <30% on DBS at 2 to 3 days	Re-testing at 5 to 10 days (profound <10%, partial 10 to 30%) and genotype sequencing	n=19 screen positive, n=14 confirmed (all partial) PPV 73.7%* PPV for profound 0%*

Abbreviations: DBS, dried blood spot; ERU, enzyme response units; MRU, microplate response units; PPV positive predictive value. *reviewer calculated

Three of the cohorts represent national/regional screening programmes, so would be expected cover the non-selected newborn population of that region. Thodi et al[6] report a large national screening pilot, but inclusion was restricted to those of Greek ethnicity so representation is uncertain.

The findings generally indicate that enzyme activity in a DBS sample has poor PPV for biotinidase deficiency. Performance data is inconsistent across studies but ranges from about 30% to 86%. This indicates that as a rough average around half of screen positives would be false.

The PPV for profound deficiency♦ is even lower. All but the Californian programme[4] indicate that fewer than 1 in 10 screen positives would have profound deficiency on confirmatory testing. A predominance of partial diagnoses and false positives could have important implications for a screening programme.

The wide range in PPV is likely influenced to some extent by prevalence. As discussed in Criterion 1, this is likely to show considerable global and ethnic variation. Higher population prevalence is likely to result in improved PPV for a screening test with fixed sensitivity and specificity (which cannot be calculated, as discussed below). Therefore it is uncertain how applicable these results would be to the UK.

However, population prevalence is unlikely to explain all of the differences in PPV. The cohorts generally followed the pattern of improved PPV with higher population prevalence, with the exception of the Californian study.[4] This had the best PPV despite the lowest prevalence. Characteristics of the screening test and its timing have likely contributed to this difference.

No evidence was identified on the optimal time to test enzyme activity to diagnose biotinidase deficiency. However, a 2010 guideline[10] noted that biotinidase activity usually increases during the first days to weeks of life. Around half of false positives are said to be due to prematurity, and a later blood spot may reduce some of these false positives. The Californian study[4] was distinct in that the screening test included a repeat DBS for those with enzyme activity in the partial range of 6 to 10 ERU. The timing of the initial or repeat tests is not given, but if it were days after the first DBS this may have removed some of the false positives and explain the PPV.

Timing of the DBS is likely to be an important factor for screening test accuracy. This is only specified for the Dutch[2] and Greek[6] programmes. The Greek programme[6] performed an early DBS at around 2 to 3 days of life, which could be expected to reduce accuracy. However, they did make the concession of delaying the screening test to 5 days for preterm

♦ It is only possible to calculate PPV for the upper cut-off of <30% (or equivalent). The cohorts do not give the number of cases with profound or partial deficiency whose initial screening test enzyme activity levels had been <10% or 10% to 30%.

infants. The other 3 studies don't mention timing of screening for preterm babies. Overall across all studies, many of the false positives could be explained by the fact that newborns may have been a week or so older when confirmatory serology was performed, by which time enzyme activity would have increased.

It is possible that repeating the DBS for those with intermediate levels, as the Californian programme,[4] could remove some false positives. Similarly lowering the enzyme activity threshold may preferentially identify those with more severe deficiency. Of note, both the Dutch[2] and Canadian[3] programmes lowered their screening test thresholds towards the end of the study period, presumably to increase test accuracy. The effects of this on PPV are not reported.

Handling and storage of the DBS sample are other factors that could have influenced screening test accuracy. The 2010 guideline[10] reports that while half of false positives may be explained by prematurity, the others could be due to mishandling of samples and possible exposure to excessive heat or humidity. These factors may have differed between studies but are not known from the study reports.

The enzyme assay method for the DBS and reference standard may also influence test accuracy. Limited detail on assay method is available in the studies, but only 2 specifically report using colorimetric assay using the artificial substrate p-aminobenzoic acid (PABA), said in the 2010 guideline[10] to be the most common method.

The reference standard also differed between studies. This could have influenced the accuracy of the diagnosis and so affect PPV of the screening test. All studies confirmed diagnosis by testing enzyme activity in repeat serology, but they variably reported whether mutation analysis or testing of parental enzyme activity was additionally required for diagnostic confirmation. In particular parental enzyme activity is said to give a better indication of whether the child could have biotinidase deficiency.[10]

Test performance data other than PPV is unknown. The lack of follow-up of screen negatives prevents calculation of test sensitivity or specificity. Wiltnik et al[2] is the only cohort that gives mention to this, stating that there were no false negatives across The Netherlands. However, it is unclear

how this was assessed. If this was assumed based on the absence of known clinical diagnoses, this may be inaccurate as some children may have had no or minimal symptoms, particularly if they had partial deficiency.

Therefore though a screening test may be expected to have high sensitivity and high negative predictive value, detecting most children with biotinidase deficiency, this cannot be known with certainty. However, it is worth noting that although PPV is generally low and specificity cannot be calculated, the false positive rate would still be expected to be low as population prevalence is low.

Summary of Findings Relevant to Criterion 4: not met.

Q3 A clear consensus on enzyme activity cut-off and other characteristics of the biotinidase deficiency screening test has not been established. Several aspects of test performance remain uncertain.

European and North American programmes demonstrate that enzyme activity in the DBS sample has a fairly low, but variable PPV, ranging from 30% to 86%. Most studies suggest that fewer than 1 in 10 screen positives will have profound deficiency. This could have important implications for a screening programme when it remains uncertain whether all those with partial deficiency need treatment.

Several factors could have contributed to the variable PPV, including differences in population prevalence and the enzyme activity threshold used and timing of the screening test (for example, performing a repeat DBS for those with enzyme activity levels in the partial range). It is not possible to calculate sensitivity and specificity as there is no follow-up of screen negatives.

Overall there is no evidence to inform whether there is an optimal cut-off that could preferentially identify those most likely to become symptomatic and need treatment; or alternatively whether repeating the screening test could remove some of the false positives.

Criterion 9 – There should be an effective intervention for patients identified through screening, with evidence that intervention at a pre-symptomatic phase leads to better outcomes for the screened individual compared with usual care. Evidence relating to wider benefits of screening, for example those relating to family members, should be taken into account where available. However, where there is no prospect of benefit for the individual screened then the screening programme should not be further considered.

Question 4 – What treatment outcomes have been reported for those with partial and profound biotinidase deficiency?

Eligibility for inclusion in the review

The 2012 UK NSC external evidence review[1] concluded that biotin is a highly effective treatment with no known adverse effects. It is also understood that early treatment for children with profound deficiency (that is, through screen detection) can prevent later symptomatic presentation and poorer outcomes.

There is more uncertainty around partial deficiency. While the vast majority of children with this type are also treated, there is practice variation around the biotin dose used. Some people with partial deficiency may have milder clinical course and the value of treating may be more limited.

This review aimed to identify RCTs, comparative cohorts or otherwise case series that reported treatment outcomes separately for those with partial and profound biotinidase deficiency.

Description of the evidence

Ten of the 73 filtered studies were initially considered relevant to this question, of which 3 were selected for full text appraisal. However, as above, there was overlap between the studies initially considered relevant to natural history.

RCTs of biotin treatment for biotinidase deficiency are likely to be unethical, and none were identified in this update search. There were also

no comparative cohorts identified that specifically aimed to assess and compare treatment outcomes according to type of deficiency or mutation.

As discussed in criterion 1, 4 cohorts/case series of diagnosed children were identified. As all children started treatment at diagnosis, these studies were considered to have most relevance to this question and are described here.

The cohorts include children identified through screening programmes. Two cohorts[2, 5] reported treatment response for children with both partial and profound deficiency, including some who discontinued treatment. Another cohort[8] described treatment response for children with profound deficiency only, some of whom discontinued treatment. One final screen cohort[3] described the effects of discontinuing treatment for three children with partial deficiency.

The findings of these studies is summarised in Table 5.

Appendix 2 contains a summary of the studies included and excluded at full text appraisal. A full summary of data extracted from each of these 4 publications is presented in the summary and appraisal of individual studies in **Error! Reference source not found.**.

Summary of findings

Table 5: Global screen cohorts reporting treatment outcomes

Study	Population/Setting	Intervention	Outcome
Wiltnik et al[2] 2016	n=50 (44 partial, 6 profound deficiency) Southwest Netherlands screening programme, 2007 to 2012.	All given biotin (dose not reported)	n=17 had mild, non-specific rashes when starting on treatment. No disease-related symptoms reported while on treatment.
Gannavarapu et al[3] 2015	n=71 (54 partial, 17 profound) Ontario, Canada screening programme, 2007 to 2014.	Starting doses reported to be 5-10mg, but maintenance doses may have varied	Outcomes reported for n=3 with partial deficiency who discontinued treatment: n=1 at age 5 years for uncertain period: neuropsychological problems, fatigue, rashes, alopecia: resolution with restarting 10mg n=1 at age 5 months for 4 to 5years: developmental, speech and hearing problems, gait instability, eczema. Outcome

Study	Population/Setting	Intervention	Outcome
			on restarting treatment not reported n=1 at age 2 years for 1.5 years: remained asymptomatic: restarted at 5mg
Jay et al[5] 2015	n=142 (120 partial, 22 profound deficiency) Screening programme, Michigan, US. Children with follow-up at the Children's Hospital of Michigan. 1988 to 2012.	Profound: 10-15mg Partial: 5mg	<u>Profound: outcomes on treatment</u> n=7 with skin problems (dose increased to 15mg; response not reported) n=5 with vison problems (myopia) **No compliance problems** <u>Partial: outcomes on treatment</u> n=20 with skin problems (dose increased to 10mg; response not reported) n=11 with developmental and educational delays n=4 with hearing loss (2 congenital and unrelated to condition; 1 in a child late starting treatment) n=1 with unrelated seizures **Non-compliance outcomes:** n=5 discontinued for >1 week and remained asymptomatic n=1 discontinued for 5 years and developed hearing loss, and skin rash and ataxia which resolved when restarting treatment
Wolf[8] 2017	n=44 with profound deficiency Screening programme, Michigan US. First laboratories to perform screening, 1984 to date.	Most started on 10mg (some 5mg, some increasing to 15-20mg in adolescence)	<u>Profound: outcomes on treatment</u> n=3 with possible sensorineural hearing loss, n=2 conductive hearing loss n=3 with seizures Normal development and education, except for isolated cases of dyslexia, ADHD, and Willian-Beuren syndrome – thought unrelated to biotinidase deficiency Several with rashes/eczema – thought unrelated to biotinidase deficiency n=5 females and 1 male experienced normal pregnancy and healthy offspring **Non-compliance outcomes:** n=1 remained asymptomatic for 20 years on suboptimal treatment (took 75 to 150µg multivitamin dose during early childhood) n=3 stopped for a few months: all experienced memory and concentration problems; 1 also experienced tremors, gastrointestinal and breathing problems; 1 also experienced hair loss and visual problems

The 4 cohorts reported treatment outcomes for individuals with biotinidase deficiency identified through regional screening programmes. The studies were not designed with the aim of comparing outcomes for

those with partial or profound deficiency, or comparing outcomes for those treated and not treated.

All children diagnosed in these screening programmes were treated. Given that adverse outcomes are only reported for a small proportion of all screen-detected individuals across cohorts, this may suggest that biotin on the whole prevents the neurological and skin problems associated with biotinidase deficiency. The Netherlands[2] and Canadian[3] cohorts in fact do not document any disease-related symptoms among treated individuals.

However, it is not possible to know whether or not those who remained asymptomatic on biotin would have developed symptoms without treatment. Differences in follow-up could also explain why adverse outcomes were reported in the US Michigan cohorts♠[5, 8] but not the other studies. The US studies have followed cases into young adulthood, whereas the Dutch[2] and Canadian[3] cohorts only cover screening experience over a short-term period of 5 to 7 years.

Where isolated adverse outcomes are reported in the Michigan cohorts[5, 8] it is difficult to know whether all of those were caused by biotinidase deficiency (for example, eczema/rashes, myopia or educational problems). The studies did not always have follow-up data on whether symptoms responded to an increase in biotin dose (which could suggest they were due to deficiency).

The Michigan[5, 8] and Canadian[3] cohorts report outcomes for a small number of individuals who discontinued treatment, but there is no consistent pattern to effect. Six with partial deficiency remained asymptomatic without treatment while 3 developed neurological and skin problems. Of the 3 people who did become symptomatic, symptoms resolved when treatment was restarted in 2 cases, but no follow-up was reported for the third. Three people with profound deficiency developed possible disease-related symptoms when stopping biotin for a few months. It is not known whether symptoms resolved when restarting

♠ It is unclear whether there could be some overlap between certain individuals in the 2 Michigan cohorts. It may be that some children initially detected by the research laboratories of the Wolf cohort also had follow-up at the Children's Hospital of Michigan and so were also included in the Jay et al cohort.

treatment. However, one person with profound deficiency remained asymptotic for 20 years while taking no biotin or extremely low dose.

In the main these studies provide anecdotal evidence which is of limited value for the purpose of informing treatment of children diagnosed through a screening programme.While all screen-detected children were treated across cohorts, the biotin dose for both partial and profound deficiency varied. It is not possible to link any of these adverse outcomes to a “suboptimal” dose or inform the best dose to prescribe for partial or profound deficiency. In any case, the small number of adverse outcome reports would limit the strength of any dose associations.

The evidence does not further understanding of which individuals with profound or partial deficiency would develop symptoms if untreated, and which would remain asymptomatic. Neither can it inform whether screening and earlier diagnosis in advance of symptom onset leads to improved outcomes compared with clinical diagnosis.

Summary of Findings Relevant to Criterion 9: not met.

Q4 Treatment outcomes for those with partial and profound biotinidase deficiency are variable.

RCTs or comparative cohorts comparing treated and untreated populations were not identified. European and North American screening programmes have treated all children with partial and profound deficiency, though dose has varied. Most have remained asymptomatic with treatment, which may suggest that biotin prevents adverse health outcomes. However, it is not known whether symptoms would have developed without treatment. Isolated adverse outcomes have been reported while on treatment, both for partial and profound, but it is not known whether all were caused by deficiency. Effects of treatment non-compliance have been inconsistent, with some individuals remaining asymptomatic (both partial and profound) without treatment and others developing symptoms.

Overall the evidence is not able to inform which screen-detected children with partial or profound deficiency would develop symptoms and need biotin treatment, or the dose to give.

Review summary

Conclusions and implications for policy

The evidence needed to support a screening programme for biotinidase deficiency in the UK is not available. The findings indicate that a change to the current policy should not be made and screening for biotinidase deficiency should not be recommended.

This review was not able to address the uncertainties raised in the last UK NSC external evidence review:[1]

1 There is no UK incidence/prevalence data available. Screening programmes and pilots from North America, Europe and the UAE find highly variable incidence, ranging from 1 in 4,508 to 1 in 31,717 for deficiency overall, and between 1 in 25,000 and 1 in 100,000 for profound biotinidase deficiency. Prevalence is likely to be influenced by ethnic and cultural variation, and it is difficult to extrapolate this data to the UK.

2 The majority of children diagnosed with partial or profound deficiency are treated with biotin. There is no data to inform the clinical course of untreated profound or partial deficiency (by enzyme activity or genotype) and explain why some people remain asymptomatic. Therefore treatment need following screening detection, particularly for partial deficiency, remains uncertain.

3 Uncertainties remain around the optimal enzyme activity threshold to use in newborn DBS screening. North American and European screening programmes show that DBS enzyme activity has a fairly low but variable PPV for biotinidase deficiency, ranging from 30% to 86%. Most studies suggest that fewer than 1 in 10 screen positives will have profound deficiency, which could have implications given the uncertainties over treatment of partial deficiency. The variation in PPV across programmes is likely explained by differences in population prevalence and characteristics of the screening test. Enzyme activity increases during the first days to weeks of life. The cohort with the best PPV (86%), despite lowest population prevalence, performed a repeat DBS for those in the partial range of enzyme activity. Other cohorts have decreased their DBS enzyme activity threshold in recent

years to improve screening test accuracy (such as changing from <30% to <20%). The optimal screening test threshold and/or timing (such as performing a later repeat DBS for those with partial levels) remain to be clarified. There is no follow-up of screen negatives, so no further test performance data (such as sensitivity and specificity) available.

4 As most children are treated at diagnosis, RCTs or comparative studies comparing treated and untreated populations are not available. Cohorts of children from North American and European screening programmes suggest that most people remain asymptomatic on biotin. The biotin dose prescribed has been variable. However, it is not known who would have remained asymptomatic without treatment. Reports of symptoms occurring while on treatment have been inconsistent, both for partial and profound deficiency. It is unclear whether all are disease-related. Similarly reported effects of treatment non-compliance, both for partial and profound deficiency, are inconsistent with some people developing symptoms and others remaining asymptomatic. The evidence is not available to inform which screen-detected children with partial or profound deficiency would develop symptoms and need biotin treatment, or the optimal dose to give. Neither can the evidence inform whether screen detection improves outcomes compared with clinical detection.

It may be difficult to obtain the evidence needed to meet the criteria for a screening programme. Children are treated at diagnosis, so study of natural history without treatment would be difficult. Also obtaining further test performance data (such as the rate of false negatives) would be difficult given that biotinidase deficiency is rare. A non-selected population screening cohort that is large enough to detect sufficient cases, while performing confirmatory re-testing for all screen negatives, is unlikely.

Nevertheless, further research would be valuable to:

1 Establish the incidence/prevalence of biotinidase deficiency in the UK.
2 Identify which levels of enzyme activity or genotypes confer increased likelihood of adverse outcomes, and which predict better outcome.
3 Better clarify the optimal threshold to use in a DBS screening test, and determine whether lower enzyme activity thresholds, or a repeat screen, could preferentially identify those who are most likely to

develop symptoms and would benefit from earlier screen detection and treatment.

4 Clarify the biotin dose to use for partial or profound deficiency.

Limitations

This was a rapid evidence review process. Searching was limited to 3 bibliographic databases and did not include grey literature sources. Literature search and first pass appraisal were undertaken by one information specialist. Second pass appraisal and study selection were then conducted by 2 analysts. Any decisions on study inclusions, queries or scope refinement were resolved in a meeting between analysts and in discussion with UK NSC evidence team as needed. The aim was to prioritise systematic reviews, RCTs and prospective controlled cohorts (according to the research questions) before moving down through the lower hierarchy of evidence.

We did not include studies that were not available in the English language, and did not review conference abstracts, conference reports or poster presentations. We were also unable to contact study authors or review non-published material.

Appendix 1 — Search strategy

Electronic databases

The search strategy included searches of the databases shown in Table . MEDLINE, MEDLINE In-Process, MEDLINE Daily, Epub Ahead of Print and Embase.

Table 6. Summary of electronic database searches and dates

Database	Platform	Searched on date	Date range of search
MEDLINE, MEDLINE In-Process, MEDLINE Daily, Epub Ahead of Print, Embase	Embase.com	12/04/17	2012 to Present
The Cochrane Library, including: - Cochrane Database of Systematic Reviews (CDSR) - Cochrane Central Register of Controlled Trials (CENTRAL) Database of Abstracts of Reviews of Effects (DARE)	Wiley Online	13/04/17	2012 to Present

Search Terms

Search terms included combinations of free text and subject headings (Medical Subject Headings [MeSH] for MEDLINE, and Emtree terms for Embase), grouped into the following categories:

- disease area
- patient group
- key questions terms

Search terms for MEDLINE, MEDLINE In-Process, MEDLINE Daily, Epub Ahead of Print and Embase for each question are shown in **Error! Reference source not found.**-10, and search terms for the Cochrane Library databases are shown in **Error! Reference source not found.**.

Question 1: What is the prevalence/incidence of biotinidase deficiency in the UK?

Table 7. Search strategy for MEDLINE, MEDLINE In-Process, MEDLINE Daily, Epub Ahead of Print and Embase

Term Group	#	Search terms	Results
Disease area	1	'biotinidase deficiency'/de	
Disease area	2	“biotinidase deficienc*”:ti,ab	
	3	#1 OR #2	
Patient group	4	'infant'/exp	
Patient group	5	'newborn'/de	
Patient group	6	'child'/exp	
Patient group	7	neonat*:ab,ti	
Patient group	8	newborn*:ab,ti	
Patient group	9	child*:ab,ti	
Patient group	10	infant*:ab,ti	
Patient group	11	#4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10	
Disease area + Patient group	12	#3 AND #11	
Epidemiology terms	13	epidemiolog*:ab,ti OR inciden*:ab,ti OR prevalen*:ab,ti	
Epidemiology terms	14	'epidemiology'/de OR 'prevalence'/de OR 'incidence'/de	
Epidemiology terms	15	#13 OR #14	
Disease area + Patient group + Epidemiology terms	16	#12 AND #15	
Date range applied	17	#16 AND (2012:py OR 2013:py OR 2014:py OR 2015:py OR 2016:py OR 2017:py)	35

Question 2: What is the natural history of profound and partial biotinidase deficiency?

Table 8. Search strategy for MEDLINE, MEDLINE In-Process, MEDLINE Daily, Epub Ahead of Print and Embase

Term Group	#	Search terms	Results
Disease area	1	'biotinidase deficiency'/de	
Disease area	2	“biotinidase deficienc*”:ti,ab	
	3	#1 OR #2	

Term Group	#	Search terms	Results
Patient group	4	'infant'/exp	
Patient group	5	'newborn'/de	
Patient group	6	'child'/exp	
Patient group	7	neonat*:ab,ti	
Patient group	8	newborn*:ab,ti	
Patient group	9	child*:ab,ti	
Patient group	10	infant*:ab,ti	
Patient group	11	#4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10	
Disease area + Patient group	12	#3 AND #11	
Natural history	13	'seizure, epilepsy and convulsion'/exp	
Natural history	14	seizure* OR convulsi*:ti,ab	
Natural history	15	'ataxia'/exp OR Ataxia:ti,ab	
Natural history	16	'visual impairment'/exp OR Vision:ti,ab OR visual:ti,ab	
Natural history	17	'hearing impairment'/exp OR Hearing:ti,ab	
Natural history	18	'rash'/exp OR 'eczema'/exp OR Rash:ti,ab OR rashes:ti,ab OR eczema:ti,ab	
Natural history	19	'alopecia'/exp OR 'hair loss'/exp OR Alopecia:ti,ab OR “hair loss”:ti,ab	
Natural history	20	'dyspnea'/de OR 'breathing difficult*':ti,ab OR 'breathing problem*':ti,ab OR dyspnea:ti,ab	
Natural history	21	'infantile hypotonia'/exp OR 'floppy baby syndrome':ti,ab OR hypotonia:ti,ab OR 'decreased muscle tone':ti,ab	
Natural history	22	#13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21	
Disease area + Patient group + Natural history	23	#12 AND #22	
Date range applied	24	#23 AND (2012:py OR 2013:py OR 2014:py OR 2015:py OR 2016:py OR 2017:py)	99

Question 3: What screening test values have been reported? Has a cut off been identified for biotinidase deficiency screening in newborns?

Table 9. Search strategy for MEDLINE, MEDLINE In-Process, MEDLINE Daily, Epub Ahead of Print and Embase

Term Group	#	Search terms	Results
Disease area	1	'biotinidase deficiency'/de	
Disease area	2	“biotinidase deficienc*”:ti,ab	
	3	#1 OR #2	
Patient group	4	'infant'/exp	
Patient group	5	'newborn'/de	

Term Group	#	Search terms	Results
Patient group	6	'child'/exp	
Patient group	7	neonat*:ab,ti	
Patient group	8	newborn*:ab,ti	
Patient group	9	child*:ab,ti	
Patient group	10	infant*:ab,ti	
Patient group	11	#4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10	
Disease area + Patient group	12	#3 AND #11	
Screening tests	13	'mass screening'/de OR 'newborn screening'/de	
Screening tests	14	screen*:ab,ti OR detect*:ab,ti OR test:ab,ti OR tests:ab,ti OR testing:ab,ti	
Screening tests	15	'diagnostic accuracy'/de	
Screening tests	16	'sensitivity and specificity'/de OR sensitivity:ab,ti OR specificity:ab,ti	
Screening tests	17	'predictive value'/de OR 'positive predictive value':ab,ti OR 'negative predictive value':ab,ti	
Screening tests	18	'likelihood ratio*':ti,ab	
Screening tests	19	'cut off':ti,ab OR 'cut-off':ti,ab OR 'cut offs':ti,ab OR 'cut-offs':ti,ab	
Screening tests	20	#13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19	
Disease area + Patient group + Screening tests	21	#12 AND #20	
Date range applied	22	#21 AND (2012:py OR 2013:py OR 2014:py OR 2015:py OR 2016:py OR 2017:py)	126

Question 4: What outcomes have been reported from treatment of those identified with profound and partial biotinidase deficiency?

Table 10. Search strategy for MEDLINE, MEDLINE In-Process, MEDLINE Daily, Epub Ahead of Print and Embase

Term Group	#	Search terms	Results
Disease area	1	'biotinidase deficiency'/de	
Disease area	2	“biotinidase deficienc*”:ti,ab	
	3	#1 OR #2	
Patient group	4	'infant'/exp	
Patient group	5	'newborn'/de	
Patient group	6	'child'/exp	
Patient group	7	neonat*:ab,ti	
Patient group	8	newborn*:ab,ti	
Patient group	9	child*:ab,ti	
Patient group	10	infant*:ab,ti	
Patient group	11	#4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10	
Disease area + Patient group	12	#3 AND #11	

Term Group	#	Search terms	Results
Treatment	13	'biotin'/de OR biotin:ti,ab	
Treatment	14	tablet*:ti,ab OR supplement*:ti,ab OR oral*:ti,ab OR treat*:ti,ab	
Treatment	15	#12 AND #14	
Disease area + Patient group + Treatment	16	#12 AND #15	
Date range applied	17	#16 AND (2012:py OR 2013:py OR 2015:py OR 2016:py OR 2017:py)	31

Table 11. Search strategy for the Cochrane Library Databases (Searched via the Wiley Online platform)

Term Group	#	Search terms	Results
	1	MeSH descriptor: [Biotinidase Deficiency] explode all trees	1
	2	"biotinidase deficienc*" (Word variations have been searched)	9
	3	#1 or #2	9

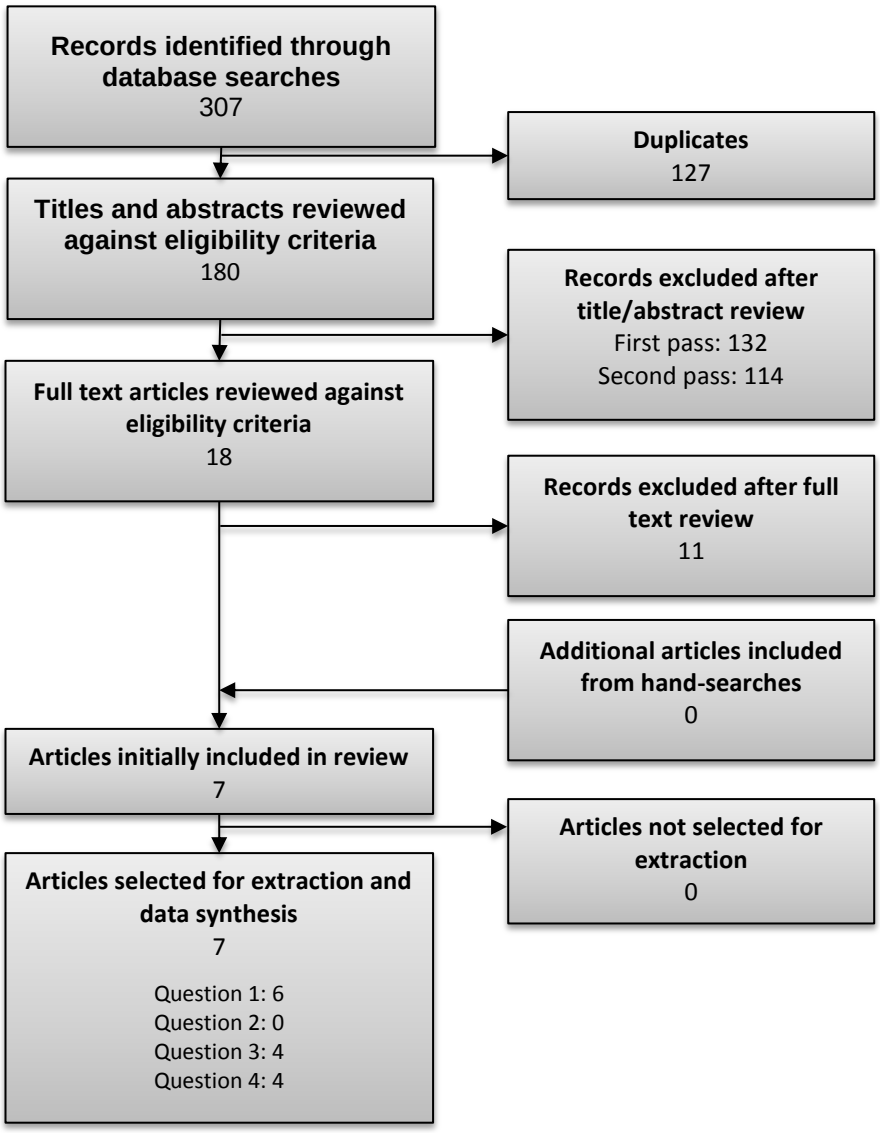
Results were imported into EndNote and de-duplicated.

Appendix 2 — Included and excluded studies

PRISMA flowchart

Figure summarises the volume of publications included and excluded at each stage of the review. Seven publications were ultimately judged to be relevant to one or more review questions and were considered for extraction. Publications that were included or excluded after the review of full text articles are summarised below.

Figure 1. Summary of publications included and excluded at each stage of the review



Publications included after review of full text articles

The 7 publications included after review of full texts are summarised in Table 12. Summary o below.

Studies were prioritised for extraction and data synthesis. It was planned *a priori* that the following approach would be taken to prioritise studies for extraction:

1 Systematic reviews and meta-analyses would be considered the highest quality of evidence if any were found. Following this, study designs would be prioritised for each question in the order listed in **Error! Reference source not found.**.
2 Studies would be prioritised if they considered a UK population, followed by studies from Western populations analogous to the UK. However, due to the limited body of evidence and lack of UK data, all studies providing relevant data were included with no restrictions to study location.

Publications not selected for extraction and data synthesis are clearly detailed in Table 12. Summary o below.

Table 12. Summary of publications included after review of full text articles, and the question(s) each publication was identified as being relevant to

Study	The condition	The test	The intervention	The screening programme	Implementation criteria	Comments
Wiltink et al 2016[2]	Q1	Q3	Q4	-	1, 4, 9	-
Gannavarapu et al 2015[3]	Q1	Q3	Q4	-	1, 4, 9	-
Jay et al 2015[5]	Q1	-	Q4	-	1, 9	-
Cowan et al 2012[4]	Q1	Q3	-	-	1, 4	-
Thodi et al 2013[6]	Q1	Q3	-	-	1, 4	-
Al Hosani et al 2013[7]	Q1	-	-	-	1	-
Wolf 2017[8]	-	-	Q4	-	9	-

Publications excluded after review of full text articles

Of the 18 publications included after the review of titles and abstracts, 11 were ultimately judged not to be relevant to this review. These publications, along with reasons for exclusion, are listed in Table 13.

Table 13. Publications excluded after review of full text articles

Reference	Reason for exclusion
Asgari A, Dehnabeh SR, Zargari M, et al. Clinical, biochemical and genetic analysis of biotinidase deficiency in Iranian population. Archives of Iranian Medicine. 2016;19(11):774-8.	No relevant data. Reports enzyme activity and associated mutation for n=8 along with parents' enzyme activity and whether consanguineous.
González Reyes EC, Castells EM, Frómeta A, et al. SUMA technology and newborn screening tests for inherited metabolic diseases in Cuba: An overview of the first 30 years. Journal of Inborn Errors of Metabolism and Screening. 2016;4.	No usable data. Describes SUMA technology relevant to Cuba and cut-off only described as "qualitative decision".
Gopalakrishnan V, Joshi K, Phadke S, et al. Newborn screening for congenital hypothyroidism, galactosemia and biotinidase deficiency in Uttar Pradesh, India. Indian Pediatrics. 2014;51(9):701-5.	Limited applicability to UK. Rural/inner city India. Shows how recall rate reduced when increasing the enzyme threshold to suit local population.
Işeri-Erten SO, Dikmen ZG, Ulusu NN. Comparison of Spectrophotometric and Fluorimetric Methods in Evaluation of Biotinidase Deficiency. Journal of Medical Biochemistry. 2016;35(2):123-9.	Case control design evaluating enzyme activity by two different assay methods.
Juan-Fita MJ, Egea-Mellado JM, González-Gallego I, et al. Expanded newborn screening in the Region of Murcia, Spain. Three-years experience. Medicina Clinica. 2012;139(13):566-71.	English language not available
Karimzadeh P, Ahmadabadi F, Jafari N, et al. Biotinidase deficiency: A reversible neurometabolic disorder (An Iranian pediatric case series). Iranian Journal of Child Neurology. 2013;7(4):47-52.	Case series n=16 clinically diagnosed. Describes presenting symptoms with

Reference	Reason for exclusion
	brief mention of improvement post-treatment but no distinction into partial or profound.
Landau YE, Waisbren SE, Chan LMA, et al. Long-term outcome of expanded newborn screening at Boston children's hospital: benefits and challenges in defining true disease. Journal of Inherited Metabolic Disease. 2017;40(2):209-18.	Case series n=45 screen-detected. Gives average developmental scores at <3 years by partial and profound, but unclear whether treated or not.
Pollak A, Kasper DC. Austrian Newborn Screening Program: A perspective of five decades. Journal of Perinatal Medicine. 2014;42(2):151-8.	Narrative discussion of screening programme, no data relevant to biotinidase deficiency.
Seoane-Mato D, Queiro-Verdes T, Atienza-Merino G, et al. Neonatal screening for biotinidase deficiency (Structured abstract). Health Technology Assessment Database: Galician Agency for Health Technology Assessment (AVALIA-T); 2014.	English language not available
Singh A, Lomash A, Pandey S, et al. Clinical, biochemical and outcome profile of biotinidase deficient patients from tertiary centre in Northern India. Journal of Clinical and Diagnostic Research. 2015;9(12):SC08-SC10.	Case series of n=10 Indian children primarily looking at the number of days' treatment needed to control seizures. No distinction into partial or profound.
Szymańska E, redzińska M, Ługowska A, et al. Outcomes of oral biotin treatment in patients with biotinidase deficiency - Twenty years follow-up. Molecular Genetics and Metabolism Reports. 2015;5:33-5.	Case series of n=22 Polish children, mostly clinically detected. Gives treatment response for symptomatic and asymptomatic individuals, but no correlation with partial or profound.

Appendix 3 — Summary and appraisal of individual studies

Data Extraction

Table 14. Studies included at full text and providing data relevant to criteria

Study reference, review questions	Study setting and population characteristics	Prevalence and incidence outcomes	Screening test performance	Treatment outcomes	Appraisal points
Wiltnik et al 2016[2] Q 1, 3, 4	Southwest Netherlands biotinidase deficiency screening programme, 2007 to 2012. n=304,982 screened by newborn dried blood spot (DBS), n=50 with confirmed diagnosis. In the whole country: n=913,965 screened, n=111 confirmed (2008 to 2012 only)	1 in 6,100 in the Southwest. Reviewer calculated: 1 in 50,830 profound 1 in 6,931 partial 1 in 8,233 for the whole Netherlands (no distinction by partial/profound).	Screen positive: biotinidase <30% on DBS tested at 5 days (range 4 to 9). Changed to <20% in July 2012. Confirmation: re-testing by colorimetric assay at 12 days (range 7 to 18) (profound <10%; partial 10-30%). Genotyping not reportedly required for confirmation. <u>Southwest:</u> n=92 screen positive, n=50 confirmed (n=6 profound, n=44 partial) Reviewer calculated: positive predictive value	Reports mild and non-specific rashes in 17/50 when referred, with no significant difference between partial and profound. All treated (dose not given). States no seizures or other disease-related symptoms, either at diagnosis or thereafter.	Nationally representative data with large sample size and incidence based on diagnostic confirmation. However, no distinction by partial or profound for national data. There may be variation from UK incidence. Cannot give additional performance data. No follow-up of screen negatives. Says no false negatives across Netherlands but unclear how that was assessed (if by clinical detection,

Study reference, review questions	Study setting and population characteristics	Prevalence and incidence outcomes	Screening test performance	Treatment outcomes	Appraisal points
			(PPV) of screening test upper threshold: 50/92=54.3% PPV for profound: 6/92=6.5% <u>Netherlands</u> n=261 screen positive, n=111 confirmed Reviewer calculated: PPV of screening test: 111/261=42.5% (No distinction by partial/profound)		some may have no/minimal symptoms). Uncertain whether storage could affect DBS results or whether reference standard is gold standard diagnosis Authors consider the predominance of partial diagnoses place an unnecessary burden, but as all were treated it is not possible to know whether or not partial would have developed symptoms.
Gannavarapu et al 2015[3] Q 1, 3, 4	Ontario, Canada biotinidase deficiency screening programme data from 5 centres, 2007 to July 2014. n=1,043,895 screened by newborn DBS, n=71 with confirmed diagnosis.	Reviewer calculated: 1 in 14,703 (71/1,043,895) 1 in 61,406 profound 1 in 19,331 partial	Screen positive: <10 enzyme response units (ERU) using Astoria Pacific Spotchek Assay on DBS; changed in Jan 2014 to 35 microplate response units (MRU). Confirmation: re-testing by serology. Most children received genotype sequencing and some parents underwent testing but not apparently required	Reports clinical course for n=3 with partial deficiency on discontinuation. n=1 discontinued age 5 followed by memory loss, fatigue, difficulty focusing, rashes, abdominal pain, slight alopecia. Resolved with increase to 10mg. Lower than average neuropsychological assessments age 8. n=1 started on 5mg and stopped treatment at 5 months to 4 to 5 years of age. At 8	Incidence data may be higher than other regions in Canada as Ontario has ethnically diverse population, and uncertain applicability to the UK. Uncertain applicability of assay method or screening test threshold to UK. Cannot give additional performance data. No follow-up of screen

Study reference, review questions	Study setting and population characteristics	Prevalence and incidence outcomes	Screening test performance	Treatment outcomes	Appraisal points
			for confirmation. n=246 screen positive, n=71 confirmed (n=17 profound, n=54 partial) Reviewer calculated: PPV of screening test upper threshold: 71/246=28.9% PPV for profound: 17/246=6.9%	years slight developmental delay, hearing difficulty, gait instability, lung infection, speech difficulty, eczema. Treatment restarted but outcome not reported. n=1 started on 2.5mg and stopped for 1.5 years at age 2 when he developed a knee infection. Remained asymptomatic during that period, then restarted 5mg with good compliance.	negatives. Follow-up for all children not reported, only for 3 who discontinued. Most starting doses are reported to be 5 to 10mg but may have varied.
Jay et al 2015[5] Q 1, 4	Retrospective review of adults with biotinidase deficiency identified since the introduction of screening in Michigan in 1988 with follow-up to 2012 at the Children's Hospital of Michigan. n=229 confirmed with deficiency during this period: n=32 profound and n=197 partial. With follow-up: n=22 profound (mean age	1 in 14,539 (across Michigan) 1 in 101,779 profound 1 in 16,533 partial	No data.	<u>Profound</u> 10 to 15mg biotin daily commenced at 1-3 weeks of age. None had hypotonia, seizures, developmental delay, optic atrophy, or hearing problems. n=7 had rashes, eczema, and contact dermatitis (biotin usually increased to 15 mg but response not documented). n=5 had vision problems including myopia and astigmatism.	Incidence data based on representative screen sample but for Michigan only which may differ from rest of the US or UK. Mostly anecdotal evidence. Hard to know if all symptoms are attributable to biotin deficiency, or how they would have developed without treatment. Some discrepancy in discussion about the number of cases. For

Study reference, review questions	Study setting and population characteristics	Prevalence and incidence outcomes	Screening test performance	Treatment outcomes	Appraisal points
	8.3 years, range 1 to 20 years) and n=120 partial (mean age 5.8 years, range 2 weeks to 29 years)			No problems with compliance. <u>Partial</u> Median biotin dose 5mg daily. Skin problems: n=20 reported including eczema, contact dermatitis, thinning hair, itching, peeling and folliculitis (biotin usually increased to 10 mg but response not documented). Developmental delays: n=4 speech/language, n=1 dyslexia, n=1 autism, n=1 delays on developmental assessment, n=1 mild cognitive deficits, n=1 special education, n=1 had to repeat a year. Hearing loss: n=2 had congenital bilateral loss not attributed to deficiency. n=1 had borderline bilateral hearing loss (dose increased to 10mg, response not reported). n=1 did not start biotin until several years of age, developed	example, conclusion discusses n=6 with hearing loss (n=1 with congenital and n=5 with sensorineural hearing loss) which differs from results discussion. Also n=11 described to have developmental delays of which the text describes only 10.

Study reference, review questions	Study setting and population characteristics	Prevalence and incidence outcomes	Screening test performance	Treatment outcomes	Appraisal points
				hearing loss with uncertain treatment response (reporting is conflicted). n=1 had seizures but this was attributed to neonatal stroke. No reports of hypotonia Compliance: 5 children missed biotin for >1 week and experienced no symptoms. n=1 discontinued for 5 years, returned to clinic with skin rash and ataxia which resolved when restarting 5mg, later increased to 10mg. Later found to have mild sensorineural hearing loss.	
Cowan et al 2012[4] Q 1, 3	California, US biotinidase deficiency screening programme July 2007 to June 2011. n=2,061,609 screened by newborn DBS, n=65 with confirmed diagnosis.	1 in 31,717 1 in 73,629 profound 1 in 55,719 partial (reviewer calculated 37/2,061,609)	Screen positive: ≤6 ERU on one DBS or 6 to 10 on two DBS using colorimetric enzyme assay. Confirmation: re-testing of enzyme activity in serum sample along with parental/ unrelated relative testing as controls if available. Genotype sequencing performed but not	No data	Large sample representative of Californian screening. However, 19/28 profound cases of Hispanic origin; may be variation in incidence compared with the UK due to ethnic differences Screening test and reference thresholds have uncertain applicability to the UK.

Study reference, review questions	Study setting and population characteristics	Prevalence and incidence outcomes	Screening test performance	Treatment outcomes	Appraisal points
			apparently required for diagnosis. n=76 screen positive: n=58 ERU ≤6, n=18 ERU 6 to 10 Confirmed: n=65, n=28 profound, n=37 partial PPV of screening test upper threshold (partial and profound): 85.5% Reviewer calculated: PPV for profound: 28/76=36.8%		No follow-up of screen negatives, no further performance data available.
Thodi et al 2013[6] Q 1, 3	Greece screening cohort (uncertain period) n=63,119 of Greek ethnic origin only screened by newborn DBS, n=14 with confirmed diagnosis.	1 in 4,508 (all partial)	Screen positive: biotinidase <30% on DBS tested at 2 to 3 days (5 days in premature infants). Confirmation: re-testing at 5 to 10 days (profound <10%; partial 10-30%) with molecular analysis n=19 screen positive, n=14 confirmed (all partial)	All were asymptomatic at diagnosis and treated with 10mg biotin, but no follow-up data available,	Relatively small screening sample of uncertain inclusion with only those of Greek ethnicity included. Unclear whether they could be representative of the population and uncertain UK applicability due to global/ethnic variation. Uncertain storage and of samples.

Study reference, review questions	Study setting and population characteristics	Prevalence and incidence outcomes	Screening test performance	Treatment outcomes	Appraisal points
			Reviewer calculated: PPV of screening test upper threshold: 14/19=73.7% PPV for profound: 0/19=0%		No additional performance data, no follow-up of screen negatives.
Al Hosani et al 2013[7] Q1	United Arab Emirates newborn screening programme, biotidinase pilot in 2010. n=25,000 screened, n=3 confirmed.	1 in 8,300 Reviewer calculated: 1 in 25,000 profound 1 in 12,500 partial (2/25,000)	Screen positive: biotinidase <30% on DBS tested at 3 days using Delfia time-resolved fluorescence assay. Confirmation: re-testing: n=1 profound (enzyme activity <10%), n=2 partial (10 to 30%) No performance data: Screen positive rate not given.	No data.	Reported 95% screening uptake of 79,464 live births in 2010. However, the pilot only included one third. Uncertain how selected and whether they are nationally representative. May also be higher incidence due to ethnic differences from the UK. Assay method may differ from the UK. Uncertain timing of storage and confirmation of samples.
Wolf 2017[8] Q4	Michigan, US, reports from the laboratories that first detected biotinidase deficiency in 1982. n=44 adults with profound deficiency	No data.	No data.	All children treated <6 weeks age, and all but a few individuals given 10mg biotin daily (a few given 5mg – no further detail). Most remained on dose; several increased to 15 or	Mostly anecdotal evidence. Hard to know if symptoms are attributable to biotin deficiency, or how they would have developed

Study reference, review questions	Study setting and population characteristics	Prevalence and incidence outcomes	Screening test performance	Treatment outcomes	Appraisal points
	identified at this centre and followed up to date. Mean age at follow-up 23.1 years, mean enzyme activity 2.7%.			20mg after puberty. <u>Developmental and educational outcomes</u> Most achieved normal development and education, with nothing remarkable. n=2 had dyslexia, one of whom had speech difficulties and hypotonia which resolved. n=1 had Willian-Beuren syndrome with developmental delays, attention problems, hypotonia and balance problems – all thought related to the syndrome. “Several” had ADHD <u>Clinical outcomes</u> n=1 bilateral sensorineural hearing loss (diagnosed with deficiency at 1 month of age). n=2 conductive hearing loss. n=1 “possibly” sensorineural but not treated. n=1 unconfirmed hearing problems. n=2 had isolated seizures, n=1 had seizures after an accident, n=1 had migraines. Few had received regular eye examinations but no problems	without treatment. Uncertain dosing and compliance for each individual. No information on partial deficiency.

Study reference, review questions	Study setting and population characteristics	Prevalence and incidence outcomes	Screening test performance	Treatment outcomes	Appraisal points
				reported. Several had rashes/eczema thought unrelated as all on biotin. n=5 females and n=1 male had children. Three mothers used 10mg during pregnancy, one took 20mg and one also took carnitine because she was deficient. No pregnancy problems reported and all children developmentally normal. <u>Compliance</u> n=4 reported non-compliance. n=1 stopped for several weeks to a month and developed tremors, gastrointestinal and breathing problems, muscle pain, memory loss. All resolved with treatment. n=1 stopped at 13 years and over several months developed emesis, lethargy, hair loss, concentration and visual problems. Resolved with treatment. n=1 cloudy thinking and moody when stopped for a short period. n=1 did not take a	

Study reference, review questions	Study setting and population characteristics	Prevalence and incidence outcomes	Screening test performance	Treatment outcomes	Appraisal points
				pharmacological dose of biotin for 20 years and was on a multivitamin containing 75 to 150µg biotin for most of early childhood and no symptoms reported. Recently restarted biotin.	

Appraisal for quality and risk of bias

Quality assessments of studies included for Criterion 4 on test validity are reported below.

Table 15. QUADAS-2 assessment of diagnostic validity studies

Domain	Risk of bias				
	Wiltink et al 2016[2]	Cowan et al 2012[4]	Gannavarapu et al 2015[3]	Thodi et al 2013[6]	Notes
Domain I: Patient selection					
Consecutive or random sample of population enrolled?	Low	Low	Low	Unclear	Thodi was a screen cohort including only a small sample and exclusively those of Greek ethnicity. Others are national/regional screening programmes with no apparent exclusions
Case-control design avoided?	Low	Low	Low	Low	
Inappropriate exclusions avoided?	Low	Low	Low	Unclear	As above
Domain II: Index test					
Index test results interpreted without knowledge of reference standard results?	Low	Low	Low	Low	Assumed on the basis that index/screen test was the indication for confirmatory testing
Threshold pre-specified?	Unclear	Low	Unclear	Low	All studies specified the threshold used but there was variation. Wiltnik and Thodi used

Domain	Risk of bias				
	Wiltink et al 2016[2]	Cowan et al 2012[4]	Gannavarapu et al 2015[3]	Thodi et al 2013[6]	Notes
					the standard 30% threshold while Gannavarapu and Cowan used ERU. Wiltnik and Gannavarpu later adapted their threshold.
Domain III: Reference standard					
Reference standard likely to correctly classify condition?	Unclear	Unclear	Unclear	Unclear	All studies report the screen test in practice and are not designed as diagnostic validity studies. In general they provide minimal information on confirmation. All have used serum enzyme activity in the reference standard, but not all report whether genotyping was required for confirmation. Only Cowan reported parent/relative testing required as standard. The gold standard for diagnosis – such as whether this should be enzyme activity, mutation analysis and parental testing – is uncertain
Reference standard results interpreted without knowledge of index test results?	Unclear	Unclear	Unclear	Unclear	Not reported by any studies
Domain IV: Test strategy flow and timing					
Appropriate interval between index test and reference standard?	Unclear	Unclear	Unclear	Unclear	Only Wiltnik and Thodi give the time interval between tests. However, the optimal interval is uncertain, particularly given that enzyme activity tends to be lower with prematurity. Only Thodi specifically reported making adjustment for prematurity in screen testing

Domain	Risk of bias				
	Wiltink et al 2016[2]	Cowan et al 2012[4]	Gannavarapu et al 2015[3]	Thodi et al 2013[6]	Notes
Did all participants receive same reference standard?	High	High	High	High	No studies follow-up screen negatives
All patients included in analysis?	High	High	High	High	As above
Applicability					
Applicable to UK screening population of interest?	Unclear	Unclear	Unclear	Unclear	General newborn population is applicable, but UK prevalence and compatibility of ethnic mix is uncertain
Applicable to UK screening test of interest?	Unclear	Unclear	Unclear	Unclear	DBS test with analysis of enzyme activity is likely to be used, but thresholds that would be used and diagnostic confirmation are uncertain
Target condition measured by reference test applicable to UK screening condition of interest?	Low	Low	Low	Low	All are assessing biotinidase deficiency

Appendix 4 – UK NSC reporting checklist for evidence summaries

All items on the UK NSC Reporting Checklist for Evidence Summaries have been addressed in this report. A summary of the checklist, along with the page or pages where each item can be found in this report, is presented in Table 2.

Table 2. UK NSC reporting checklist for evidence summaries

	Section	Item	Page no.
1.	TITLE AND SUMMARIES		
1.1	Title sheet	Identify the review as a UK NSC evidence summary.	Title page
1.2	Plain English summary	Plain English description of the executive summary.	5
1.3	Executive summary	Structured overview of the whole report. To include: the purpose/aim of the review; background; previous recommendations; findings and gaps in the evidence; recommendations on the screening that can or cannot be made on the basis of the review.	8
2.	INTRODUCTION AND APPROACH		
2.1	Background and objectives	Background – Current policy context and rationale for the current review – for example, reference to details of previous reviews, basis for current recommendation, recommendations made, gaps identified, drivers for new reviews	13
		Objectives – What are the questions the current evidence summary intends to answer? – statement of the key questions for the current evidence summary,	15

	Section	Item	Page no.
		criteria they address, and number of studies included per question, description of the overall results of the literature search. Method – briefly outline the rapid review methods used.	17
2.2	Eligibility for inclusion in the review	State all criteria for inclusion and exclusion of studies to the review clearly (PICO, dates, language, study type, publication type, publication status etc.) To be decided *a priori*.	18
2.3	Appraisal for quality	Details of tool/checklist used to assess quality	20
3.	SEARCH STRATEGY AND STUDY SELECTION (FOR EACH KEY QUESTION)		
3.1	Databases/ sources searched	Give details of all databases searched (including platform/interface and coverage dates) and date of final search.	20
3.2	Search strategy and results	Present the full search strategy for at least one database (usually a version of Medline), including limits and search filters if used. Provide details of the total number of (results from each database searched), number of duplicates removed, and the final number of unique records to consider for inclusion.	41
3.3	Study selection	State the process for selecting studies – inclusion and exclusion criteria, number of studies screened by title/abstract and full text, number of reviewers, any cross checking carried out.	17
4.	STUDY LEVEL REPORTING OF RESULTS (FOR EACH KEY QUESTION)		
4.1	Study level reporting, results and risk of bias	For each study, produce a table that includes the full citation and a summary of the data relevant to the question (for example, study size, PICO, follow-up	Full extraction: 53 Summary tables: 22, 29, 35

	Section	Item	Page no.
	assessment	period, outcomes reported, statistical analyses etc.). Provide a simple summary of key measures, effect estimates and confidence intervals for each study where available. For each study, present the results of any assessment of quality/risk of bias.	
4.2	Additional analyses	Describe additional analyses (for example, sensitivity, specificity, PPV, etc.) carried out by the reviewer.	Full extraction: 51 Summary tables: 22, 28, 34 (PPV calculations within table)
5.	QUESTION LEVEL SYNTHESIS		
5.1	Description of the evidence	For each question, give numbers of studies screened, assessed for eligibility, and included in the review, with summary reasons for exclusion.	21, 25, 27, 33
5.2	Combining and presenting the findings	Provide a balanced discussion of the body of evidence which avoids over reliance on one study or set of studies. Consideration of four components should inform the reviewer's judgement on whether the criterion is 'met', 'not met' or 'uncertain': quantity; quality; applicability and consistency.	22, 28, 34
5.3	Summary of findings	Provide a description of the evidence reviewed and included for each question, with reference to their eligibility for inclusion. Summarise the main findings including the quality/risk of bias issues for each question. Have the criteria addressed been 'met', 'not met' or 'uncertain'?	26, 32, 37
6.	REVIEW SUMMARY		
6.1	Conclusions and implications for	Do findings indicate whether screening should be recommended?	38

	Section	Item	Page no.
	policy	Is further work warranted? Are there gaps in the evidence highlighted by the review?	
6.2	Limitations	Discuss limitations of the available evidence and of the review methodology if relevant.	40

References

1. NSC. Newborn screening for biotinidase deficiency: external review against programme appraisal criteria for the UK National Screening Committee (UK NSC). London: UK National Screening Committee, 2012. Available from: https://legacyscreening.phe.org.uk/policydb_download.php?doc=264.
2. Wiltink RC, Kruijshaar ME, Van Minkelen R, et al. Neonatal screening for profound biotinidase deficiency in the Netherlands: Consequences and considerations. European Journal of Human Genetics. 2016;24(10):1424-9.
3. Gannavarapu S, Prasad C, DiRaimo J, et al. Biotinidase deficiency: Spectrum of molecular, enzymatic and clinical information from newborn screening Ontario, Canada (2007-2014). Molecular Genetics and Metabolism. 2015;116(3):146-51.
4. Cowan TM, Kazerouni NN, Dharajiya N, et al. Increased incidence of profound biotinidase deficiency among Hispanic newborns in California. Molecular Genetics and Metabolism. 2012;106(4):485-7.
5. Jay AM, Conway RL, Feldman GL, et al. Outcomes of individuals with profound and partial biotinidase deficiency ascertained by newborn screening in Michigan over 25 years. Genetics in Medicine. 2015;17(3):205-9.
6. Thodi G, Schulpis KH, Molou E, et al. High incidence of partial biotinidase deficiency cases in newborns of Greek origin. Gene. 2013;524(2):361-2.
7. Al Hosani H, Salah M, Osman HM, et al. Expanding the comprehensive national neonatal screening programme in the United Arab Emirates from 1995 to 2011. Eastern Mediterranean Health Journal. 2013;20(1):17-23.
8. Wolf B. Successful outcomes of older adolescents and adults with profound biotinidase deficiency identified by newborn screening. Genetics in Medicine. 2017;19(4):396-402.
9. Szymańska E, redzińska M, Ługowska A, et al. Outcomes of oral biotin treatment in patients with biotinidase deficiency - Twenty years follow-up. Molecular Genetics and Metabolism Reports. 2015;5:33-5.
10. Cowan TM, Blitzer MG, Wolf B, et al. Technical standards and guidelines for the diagnosis of biotinidase deficiency. Genet Med. 2010;12(7):464-70.