UK National Screening Committee

Antenatal screening for cystic fibrosis

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

Version: Final

Author: Bazian

Date: 13th February 2019

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 5, Wellington House, 133-155 Waterloo Road, London, SE1 8UG
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 2019

Contents

Plain English summary

Cystic fibrosis (CF) is an inherited condition that causes thick sticky mucus to build up in the lungs and digestive system. This causes problems with digesting food as well as lung infections. A faulty gene called the cystic fibrosis transmembrane conductance regulator (*CFTR*) causes CF.

People born with CF have inherited 2 copies of a faulty *CFTR* gene, 1 from each of their parents. If parents are carriers, they have 1 faulty gene copy and 1 healthy gene copy, so they do not have CF themselves. But, there is a 1 in 4 chance that carrier parents will pass their faulty genes to their baby who will develop CF. About 1 in 25 white European people are carriers.

Currently babies in the UK are screened for CF as part of the newborn screening programme. The purpose of newborn screening is to diagnose the baby early so they can receive the care that they need.

This review aims to see if there is evidence to support the introduction of a screening programme for CF in pregnancy in the UK.

Pregnancy screening would involve testing both parents to see if they are carriers of a faulty *CFTR* gene. If both parents are found to be carriers they can be offered further testing to see if the baby has inherited a faulty gene from each parent and will have CF. As there is no cure for CF, the purpose of pregnancy screening is to allow carrier parents to make fully-informed pregnancy decisions.

The review found that:

- around 3 in 20,000 people in the UK have CF and there are around 6 new cases in 20,000 births each year
- there are many types of faults with the *CFTR* gene but we are not able to link the faults to the seriousness of disease. So at the moment it is not possible to give information to pregnant couples about how their baby will be affected by the disease
- it is unclear which faulty genes a pregnancy screening programme should look for
- it is also unclear if pregnancy screening is acceptable to the general population and to those affected by CF in the UK

These uncertainties suggest that further research is needed. There is not enough evidence to recommend a pregnancy screening programme for CF in the UK.

Executive summary

Purpose of the review

This review aimed to see whether the evidence is available to support a population-wide antenatal screening programme for cystic fibrosis (CF).

Background

CF is an autosomal (non-sex-linked) recessive condition caused by disease-causing variants (mutations) of the cystic fibrosis transmembrane conductance regulator (*CFTR*) gene. It is estimated that around 1 in 25 people of white European ethnicity carry one copy of a disease-causing variant. If both parents carry one copy there is a 1 in 4 chance that the baby will inherit one disease-causing variant from each parent and so develop CF.

The *CFTR* gene codes for a protein that transports chloride across the membrane of epithelial cells. This in turn regulates the flow of sodium, bicarbonate, potassium and water. If the transporter protein does not function correctly it results in thick mucus build up in the lungs, digestive system and other exocrine organs. Poor lung function, chest infections, pancreatic dysfunction and nutritional deficiency are common. Males with CF are almost always infertile.

However, there are more than 2000 *CFTR* gene variants currently recognised and severity of disease can vary widely. Variants are increasingly being grouped into 5 classes depending on the effect that the gene variant has on production or processing of the protein product. Broadly, variants in classes 1-3 result in no functioning protein in the membrane, while variants in classes 4 and 5 may allow some residual protein function.

Current CF screening programmes in the UK

Since 2007 screening for CF in the UK has been carried out as part of the newborn blood spot (NBS) screening programme. The purpose of newborn screening is to allow for early diagnosis and treatment.
There is no curative treatment for CF but life expectancy continues to improve. The median predicted life expectancy for those born in 2012-16 is 47 years[1] compared with 35 years in 2007 and 39 years in 2008.[2] This is mostly due to treatment advances. Previously care was mostly supportive but new treatments are being developed that target the functional defect in people with specific variants, such as the drug ivacaftor.

Carrier testing is also available for blood relatives of people with a diagnosis of CF and their partners (cascade screening). Couples who are carriers may be offered testing (such as chorionic villus biopsy or amniocentesis) during pregnancy, to identify if the baby has the condition. A small proportion of CF carrier babies are also detected incidentally through the newborn screening programme. However, population-based or universal antenatal carrier screening is not currently performed in the UK.

Focus of the review

The current review aimed to see whether the evidence is available to support population-based antenatal screening for CF.

Such a programme would either screen both parents at the same time (couple screening) or sequentially, where the second parent was screened only if the first parent was found to be a CF carrier (stepwise or sequential screening). If both parents were carriers, the couple could be offered antenatal testing (amniocentesis or chorionic villus sampling) to see if the baby carried 2 disease-causing CF variants. The purpose would be to support informed pregnancy decision making.

In order to support this, the review needed to establish whether there is sufficient understanding in several areas.

The review addressed 4 key questions:

1. What is the UK prevalence of CF and CF carrier status among the general population, by genotype and by ethnicity? Has prevalence changed over time? (Criterion 1)
2. What are the genotype-phenotype associations in cystic fibrosis patients, including their clinical prognosis? (Criterion 1)
3. What genotypes/variants are covered by commercially available antenatal CF screening tests in the UK? What is the clinical sensitivity of these tests for predicting CF in the fetus/newborn? (Criteria 4 and 8)
4. Is an antenatal screening programme acceptable to people in the UK, specifically to pregnant women and their partners, to people with CF carrier status, and people affected by CF (patients or family members)?(Criterion 12)

A rapid review search was undertaken for questions 1, 3 and 4. The search for these questions was conducted in April 2018 for studies published from 2000 onwards (following publication of the last health technology appraisal on this issue, as below).

If antenatal CF screening is to be used to guide pregnancy decisions based on genotype, there needs to be a clear association between genotype and clinical outcome. Therefore a full systematic review approach was undertaken for question 2 on the genotype-phenotype

association. The search was conducted in May 2018 with no date limit and citations of included studies were hand-searched.

Recommendation under review

The UK National Screening Committee (NSC) does not currently recommend universal antenatal screening for CF. This policy was last considered in July 2006, the same time at which newborn screening for CF was reviewed. Newborn CF screening was implemented across England and Wales in 2007. In Northern Ireland the protocol for the existing newborn CF screening programme was amended to include CF mutation analysis in 2009. Given the emphasis on newborn screening the Committee decided not to undertake a review of antenatal screening at that point.

These policy decisions were primarily informed by a 1999 Health Technology Appraisal (HTA).[3] This HTA had reviewed antenatal screening alongside alternative screening options of preconception, population, newborn and cascade screening.

Key findings from the 1999 review were that:

- prevalence of CF and CF variants varies across regional and ethnic populations
- there was practical antenatal screening experience from 11 screening pilots (including 5 from the UK) which found that the screening uptake rate was 70%, 89% of carrier couples identified opted for antenatal diagnosis, and nearly all pregnancies where the fetus was found to carry 2 CF-causing variants were terminated
- it was not possible to predict the clinical course of disease even with variants known to be associated with severe phenotype due to potential confounders like treatment availability
- a negative carrier status cannot exclude the possibility that the fetus may be born with CF as there are so many disease-causing variants
- screening may be associated with adverse psychological and emotional effects

The 4 key questions in this review aimed to address these broad areas and see whether there is new literature to inform these gaps in the evidence.

Findings and gaps in the evidence of this review

The review found that the volume, quality, applicability and direction of evidence published do not comprehensively answer these questions, leaving several remaining uncertainties:

1. Data is available from the UK CF registry which shows that in 2016 CF affected 1 in 6276 people in the UK or 1.59 per 10,000 of the population. The incidence in 2016 was 1 in 3137 live births or 3.19 per 10,000 births per year. Birth incidence increased in 2007, the timing of introduction of NBS for CF. Since 2007 there has been no clear change in incidence, but

prevalence has continued to rise since the Millennium. This suggests that survival may be improving. There has been little change in genotype prevalence over the years. Variant c.1521_1523delCTT (hereafter referred to by the legacy name F508del) is by far the most common variant carried by 90% of people with CF. Around 50% carry 2 copies of this variant (homozygotes). The UK CF registry covers 99% of people with CF seen in clinics across the UK, so is likely to give a true reflection of prevalence and incidence. **Criterion 1 – prevalence and incidence – met**. There was no data on prevalence or incidence by ethnicity or on CF carrier prevalence.

2. There is consistent evidence from 15 large studies that gene variant class 1 to 5 is linked with phenotype in CF. F508del homozygotes and compound heterozygotes who carry 2 copies of a class 1-3 variant are likely to have pancreatic insufficiency and poor survival outlook. People who carry at least one class 4 or 5 variant are likely to have milder disease course with lower rates of pancreatic insufficiency and longer survival. There were similar but less consistent associations with lung function and age at diagnosis. However, across studies phenotype was highly variable for people with the same genotype or with variants in the same functional class. One study looked at the ability of genotype class to predict age at death. It found that while most people who die before age 30 years carry 2 severe class 1-3 variants, a third with these genotypes live beyond this age. Similarly around a third of people with at least one mild class 4 or 5 variants would still die before 30. Studies generally found that around 90-100% of people with 2 class 1-3 variants including F508del homozygotes had pancreatic insufficiency, and were usually diagnosed prior to the age of 2 years. However, between 25% and 75% of people with at least one class 4 or 5 variant also had pancreatic insufficiency, and although diagnosis was usually later, it varied from childhood to adulthood. Therefore it would not be possible to accurately predict individual disease course with any certainty based on genotype alone.

   There are also several limitations to the evidence. Most studies are based on registry data and genotype or classification information was not available for typically half of the registry population. Therefore results may not represent the CF population as a whole. Few studies adjusted for treatment or care received and other confounding variables, increasing the risk of bias. Furthermore most cohorts date from over 20 years ago and may not be applicable today because treatment advances may have considerably altered prognosis. Finally, only a few potentially disease-causing CF variants have been widely studied, classified or included in prior antenatal screening panels. The phenotypic effects of many rare variants are unknown.

   Overall, there is evidence of an association between genotype and phenotype. However, due to the variability in outcomes for individuals, risk of bias across studies (particularly relating to lack of genotyping and confounding), limited applicability to care today, and uncertain effects of rare variants, there is insufficient evidence to reliably predict the genotype-phenotype association. **Criterion 1 – genotype-phenotype association – not met.** This degree of uncertainty is considered a reasonable price to pay in newborn screening as more babies will benefit than be harmed from screening. However, in antenatal testing where the option is to continue or terminate the pregnancy, a much higher degree of certainty is needed. Furthermore, in the newborn programme, mutation analysis is a second step only carried out for infants with immunoreactive trypsinogen levels above the cut-off (on 2 assays).
3. No studies have been published investigating antenatal screening in the UK since 2000. Only a single screening pilot has been conducted in Victoria, Australia. This study screened 3200 individuals and detected 106 carriers with a carrier frequency of 1 in 30. Subsequent screening

of their partners (sequential testing) identified 6 pregnant carrier couples, all of whom accepted diagnostic testing. The 2 positive pregnancies (positive predictive value 33%) were both terminated – consistent with findings of pre-2000 screening pilots. There was no follow-up of screen-negatives so further test accuracy data was not available. This study also had limited applicability to the UK as it was a pay-for service, included preconception screening and tested for variants prevalent in the local population (not all of which are common in the UK). Pre-2000 UK pilots had also differed in the variants they tested for and the background literature indicates that there is as yet no well-established variant panel that could be used in an antenatal screening test for CF in the UK. **Criteria 4 and 8 – not met.**

4. No studies have assessed views on universal antenatal CF screening among the UK population. A sample of non-UK literature identified by the search included views of people taking part in the post-2000 Australian screening pilot. This generally indicated a lack of understanding about CF screening, for example, believing if you received a negative test result you were definitely not a carrier of any CF disease-causing variants; high levels of anxiety about antenatal diagnosis among couples who screened positive; and grief and regret over termination decisions. An additional Belgian study questioning views of people affected by CF (majority Catholic) found concerns that it would detract resources from CF and increase termination rates. These studies do not represent all of the international literature on screening views and are culturally-specific so cannot be generalised to the UK. On the basis of no UK evidence this criterion is not met. **Criteria 12 – not met.**

Recommendations on screening

The findings indicate that the current policy not to perform population-wide antenatal screening for CF should not be reversed at the current time.

Limitations

The search strategy was built on a protocol developed *a priori* for each of the 4 key questions. Searching was limited to 3 literature databases (4 for question 2 on genotype-phenotype association) and did not include grey literature resources for questions 3 and 4. Studies only available in non-English language, editorials, abstracts, conference reports or poster presentations were not included. The reviewers were unable to contact study authors or review non-published material. The systematic review on genotype-phenotype association has not analysed the effect of complex alleles (more than one variant on the same allele) or the influence of environmental factors or genes other than *CFTR* that may mediate the genotype-phenotype association.

Evidence uncertainties

Further research may help to address the uncertainties around each of the 4 key questions:

1. Information on the carrier prevalence of CF variants among the general UK population, overall and by ethnicity. Information on the prevalence and incidence of CF by ethnicity.
2. Improved understanding of the phenotypic effects of rarer CF variants, and of the influence that modifier genes (other than *CFTR*), complex alleles (more than one disease-causing variant on the same allele) and environmental factors may have on genotype-phenotype relationships.
3. To establish a panel of variants that could be used in antenatal screening in the UK and to conduct further antenatal screening pilots in the UK that use these variants. Such studies would benefit from conducting longer term follow-up and surveillance of all screen-negatives to give an indication of clinical sensitivity, specificity, positive and negative predictive values of the test.
4. Study of the whether a population-wide antenatal screening programme is acceptable in the UK, to the population in general, to carriers and to people affected by CF.

Introduction and approach

Background

Cystic fibrosis (CF) is an autosomal (non-sex-linked) recessive condition caused by disease-causing variants (mutations) of the cystic fibrosis transmembrane conductance regulator (*CFTR*) gene. It is the most common hereditary single gene condition in the UK.[4] If both parents carry one disease-causing variant of the *CFTR* gene there is a 1 in 4 chance that a baby will inherit abnormal variants from each parent and so develop CF. Birth prevalence among white Europeans has long been estimated at around 1 in 2500 births, and 1 in 25 people are carriers (carrying one disease-causing variant).

The *CFTR* gene is located on chromosome 7 and codes for a transporter protein found in the membrane of epithelial cells lining the secretary exocrine glands. The CFTR protein is made up of two membrane-spanning domains that form a chloride channel. The transport of chloride in turn regulates the flow of sodium, bicarbonate, potassium and water across the cell membrane.[5, 6]

Abnormalities in the CFTR transporter protein cause thickened sticky mucus secretions in the lungs, digestive system and other exocrine organs resulting in multi-systemic symptoms. Poor lung function, chest infections, pancreatic dysfunction, nutritional deficiency and low bone mineral density are common in people with CF. Males are almost always infertile due to absence or blockage of the vas deferens (which transports sperm to the urethra).

Life expectancy for people with CF is reduced but is consistently improving. The median predicted life expectancy for those born in 2012-16 is 47 years[1] compared to 35 years in 2007 and 39 years in 2008.[2] This is largely due to improvements in treatment. Until recently CF treatment was mostly supportive, but newer specialised treatments are being developed that target the functional defect in people with specific variants, such as the drug ivacaftor. Ivacaftor has been licensed for the treatment of class 3 variants (see below), in the UK since 2012.

Disease-causing variants (gene mutations)

The Cystic Fibrosis Mutation Database (CFTR1) collects international data on individual *CFTR* gene variants and now documents over 2000.[7] The companion project Clinical

and Functional TRanslation of CFTR (CFTR2) has currently detailed 374 of these variants, 312 of which are believed to cause classic CF symptoms with the reminder of less certain clinical consequence.[8] However, the vast majority of variants are rare and it is estimated that only 20 have a frequency above 0.1% worldwide.[5]

Around half of people with CF will be homozygous for the most common disease-causing variant c.1521_1523delCTT (hereafter referred to by the legacy name F508del)*. The remainder will mostly be compound heterozygotes carrying 2 different CF variants, usually F508del in combination with another variant.

Various classification systems have been used in the past, including those that classify according to whether the variants produce "classical" multi-systemic disease or whether they cause "non-classical" single-organ disease or CFTR-related disorders.[9]

Currently the most widely used system classifies variants into 5 groups according to the functional effect they have on the CFTR protein: [5, 6]

1. Protein production variants (Class 1) – cause little or no CFTR protein to be produced so it is absent from the membrane
2. Protein processing variants (Class 2) – affect how the CFTR protein is processed within the cell and transported to the membrane (the most common variant F508del is typical of this class)
3. Regulation/gating variants (Class 3) – protein is present in the membrane but ion transport through the channel is impaired (G551D is typical)
4. Conduction variants (Class 4) – channel conductance is impaired but there is still residual function
5. Reduced production or processing variants (Class 5) – CFTR is present in the membrane but in reduced quantity

In general class 1 to 3 variants result in minimal functioning CFTR protein and so would be expected to cause severe disease. Class 4 to 5 variants, where some CFTR function is maintained, may confer milder disease course, even if present with another severe variant.[10]

However, the possible modulating effect of one variant upon the other, the presence of complex alleles (where there is more than one disease-causing variant on the same allele), other genes acting as modifiers and environmental factors may all influence the phenotype. Therefore predicting phenotype from genotype may be challenging.

* Throughout this report *CFTR* variants have been referred to by their legacy names, which have been used in all cited literature. For the complete list of corresponding Human Genome Variation Society (HGVS) names, see Table 12.

Screening

In the UK screening for CF is carried out as part of the newborn blood spot (NBS) screening programme. This involves measuring levels of the enzyme immunoreactive trypsinogen (IRT) in the newborn, which is elevated due to pancreatic dysfunction. Infants with IRT levels above the cut-off level (on 2 assays) are tested for the 4 most common variants in the UK that tend to be associated with severe phenotype (CF4 panel: F508del, G542X, G551D, 621+1G→T).[11] Sweat testing for raised salt levels may also be carried out to verify diagnoses. The purpose of newborn screening is to allow for early diagnosis and treatment.

Carrier testing is also currently available for blood relatives of people with a diagnosis of CF and their partners (cascade screening). Couples who are carriers may be offered testing (such as chorionic villus biopsy or amniocentesis) during pregnancy, to identify if the baby has the condition. A small proportion of CF carrier babies are also detected incidentally through the newborn screening programme. However, population-based antenatal carrier screening is not currently performed in the UK.

If a fetus is found to carry 2 causative CF variants there is no treatment available. The purpose of antenatal screening would be to provide parents with comprehensive information so that they can make an informed reproductive choice whether to continue with or terminate the pregnancy. Therefore a clear understanding is needed whether a particular genotype could reliably predict the clinical outcome (phenotype) in any individual.

Current policy context and previous reviews

The UK NSC does not currently recommend universal antenatal screening for CF. This policy was last considered in July 2006, the same time at which newborn screening for CF was reviewed. Newborn CF screening was implemented across England and Wales in 2007. In Northern Ireland the protocol for the existing newborn CF screening programme was amended to include CF mutation analysis in 2009. Given the emphasis on newborn screening the Committee decided not to undertake a review of antenatal screening at that point.

These policy decisions were primarily informed by a 1999 Health Technology Appraisal (HTA).[3] This HTA had reviewed antenatal screening alongside alternative screening options of preconception, population, newborn and cascade screening.

Key findings from this review were that:

- prevalence of CF and of different variants has been shown to differ across different regional and ethnic populations
- antenatal screening appeared to be practical and feasible, following publication of 11 screening pilots of couple or stepwise screening (5 conducted in the UK) which showed:
    - overall screening uptake of around 70%
    - subsequent antenatal diagnosis in carrier couples was 89%
    - diagnosis of a fetus carrying 2 CF variants resulted in parents opting for termination in all but one pregnancy
- as there are many disease-causing CF variants, a negative variant test (in one or both parents) would not exclude the possibility that the parents were carriers and therefore that the baby might be affected
- it was difficult to predict the clinical course of disease even with variants associated with severe phenotype (homozygous or in combination) due to potential confounders like treatment availability. While pancreatic function is established as a discriminatory clinical feature, the association with genotype was unclear
- screening is associated with various risks, including psychological and emotional:
    - many couples with negative results may falsely believe they have no risk of having a child affected by CF
    - some people experience anxiety as a result of the screening process
    - people identified as carriers may experience stigmatisation
    - prenatal diagnosis (amniocentesis or chorionic villus sampling [CVS]) carries risk of miscarriage that it is difficult to quantify
    - views on antenatal screening among people affected by CF are rarely obtained; past surveys found preference or acceptance of preconception and newborn screening but only half found termination of an affected pregnancy acceptable
- new and improved treatments were expected to improve prognosis for people affected by CF

Murray et al[3] considered at the time that “antenatal screening should be offered routinely to women and their partners in all maternity units” as this seemed the most practical approach. However, they also considered that there was a large body of indirect evidence that early diagnosis through newborn screening could improve long-term outcomes. As such Murray et al[3] also recommended that “each purchasing health authority could consider providing neonatal CF screening, either in combination with antenatal screening or alone.”

In 2007, newborn screening for CF was implemented across the UK to facilitate earlier diagnosis and treatment of individuals with CF. This changed the tenor for antenatal screening for CF and it has not been reviewed since.

Objectives

The current review aims to review and summarise the evidence on universal antenatal CF screening published since the 1999 HTA. It aims to see whether new evidence is available to suggest that the current policy not to offer antenatal screening for CF should be reconsidered.

Four questions will be addressed to cover the key issues identified by the 1999 HTA.[3] These questions are outlined in Table 1.

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

	Criterion	Key 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 of CF and CF carrier status: b) by genotype c) does it vary by ethnicity d) has it changed over time	UK CF Registry annual reports with one additional cohort
		Q2: What are the genotype-phenotype associations in cystic fibrosis patients, including their clinical prognosis?	15 studies
	THE TEST		
4 **8**	There should be a simple, safe, precise and validated screening test. If the test is for a particular mutation or set of genetic variants the method for their selection and the means through which these will be kept under review in the programme should be clearly set out.	Q3: a) to describe the genotypes/mutations covered by commercially available tests for antenatal CF screening in the UK, which have been tested in published research b) to estimate the clinical sensitivity and specificity of these tests and estimate their positive and negative predictive values	One antenatal screening pilot
	THE SCREENING PROGRAMME		
12	There should be evidence that the complete screening programme (test, diagnostic procedures, treatment/ intervention) is clinically, socially and ethically acceptable to health professionals and the public.	Q4: Is an antenatal screening programme acceptable to people in the UK: a) pregnant women and their partners b) individuals with CF carrier status c) individuals with CF	0 UK studies

Methods

The current review was conducted by Bazian, in keeping with the UK NSC evidence review process. The review was conducted in a two-phased approach.

The first phase was a rapid review to review and summarise the body of evidence addressing the 3 questions on prevalence, test accuracy and acceptability. Database searches for these 3 questions were conducted on 13th April 2018.

The second phase involved a systematic review to summarise the available literature on the genotype-phenotype correlation. Database searches for the systematic review were conducted on 11th May 2018.

Eligibility for inclusion in the rapid review

A systematic literature search of MEDLINE and Embase databases (Embase.com) and The Cochrane Library (Wiley Online) was performed for studies published between January 2000 and April 2018. The full search strategy is presented in 0.

The search yielded 1318 references meeting the search terms of the rapid review questions. These studies were further filtered at title and abstract level by one information specialist, and 98 studies considered potentially relevant to the 3 questions were selected at first sift.

Each of these abstracts was reviewed against the inclusion/exclusion criteria (as outlined in Table 2) by the main reviewer. Where applicability of inclusion was unclear, the article was included at this stage in order to ensure that all potentially relevant studies were captured.

27 studies were selected for full text appraisal, in addition to online UK CF registry data. Each full text article was then reviewed against the inclusion/exclusion criteria by the main reviewer, who determined whether the article was relevant to one or more of the review questions. A second senior reviewer provided input in cases of uncertainty.

0contains a full PRISMA flow diagram (
), along with a table of the included publications and details of the questions these publications were relevant to (Table 21. Summary o). Exclusions at full text with reason for exclusion are listed by question in Table 22.

Eligibility for inclusion in the systematic review

The scope of the systematic review on genotype-phenotype association was discussed between Bazian, the UK NSC evidence team and an external topic expert[†]. Following selection of key outcomes and finalisation of the scope, the search strategy was agreed between these members. A systematic literature search was then performed in MEDLINE and EMBASE databases (Embase.com), The Cochrane Library (Wiley Online) and Scopus on 11th May 2018. No date limits or study design filters were applied. The full search strategy is presented in 0.

The search yielded 9238 references relevant to genotype-phenotype association. These studies were initially filtered at title and abstract level by an information specialist, and 841 studies were considered potentially relevant to the question at first sift.

Each of these abstracts was reviewed against the inclusion/exclusion criteria (as outlined in Table 2) by the main reviewer. Broad sifting decisions and exclusions at abstract level were agreed in discussion with a second senior reviewer and with the topic expert, who provided guidance on the clinical outcomes that were of greatest relevance for assessment. Exclusions at abstract are further described in question 2, Criterion 1. Where applicability of inclusion was unclear from the abstract, the article was acquired at full text in order to ensure that all potentially relevant studies were captured.

76 studies were selected for full text appraisal. These studies were reviewed by the main reviewer and potential inclusions and exclusions were discussed with a second senior review. The external topic expert reviewed the list of inclusions and exclusions to check whether there were any important omissions. Citations of included articles were also hand-searched but no relevant studies on genotype-phenotype association studies were identified.

Further information on the evidence selection process is presented in question 2, Criterion 1 in the report below. 0contains a full PRISMA flow diagram (
). A list of studies excluded at full text with accompanying rationale is given in Table 22.

Due to the heterogeneity of studies in terms of their included populations, and methods of genotype comparison and outcome assessment, the decision was made not to perform

[†] The authors thank the contribution from Professor Kevin Southern, Professor of Child Health at the University of Liverpool.

meta-analysis. The findings of the studies have been discussed narratively to show the range of results for different outcomes and the overall direction of effect.

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

Key question	Inclusion criteria							Exclusion criteria
	Population	**Target condition**	**Intervention/Test**	**Reference Standard**	**Comparator**	**Outcome**	**Study type**	
1	Non-selected general UK population samples or samples of pregnant women, including by ethnicity	CF or CF carrier status, including by specific genotype	NA	NA	NA	Incidence or prevalence data by population, including change over time.	Surveillance reports/registry data. Cross-sectional or cohort studies representative of population sample (e.g. by ethnicity). Systematic reviews of these studies	High-risk samples (e.g. those with family history). Non-UK. Studies with sample size <500. Conference abstracts, non-English language studies.
2	People with CF – either grouped by genotype or phenotype depending on the study design	Diagnosed CF	any supportive treatment or disease-specific treatment given, or NA depending on the study	NA	People with other genotype(s) or alternatively people with and without phenotype if case control	survival, age at diagnosis, lung function (e.g. FEV1), quality of life, pancreatic sufficiency, treatment response	Prospective or retrospective cohort studies, case control studies, cross sectional studies, systematic reviews of these studies	Case reports and case series. Conference abstracts, editorials or non-systematic reviews. Non-English language studies. More methodological detail provided in individual section.

3	Non-selected pregnant women and their partners not known to be at risk of CF	CF carrier status in the couple CF in the fetus or newborn	Any antenatal screening test commercially available in the UK testing for carriage of disease-causing variants. Couple or stepwise screening.	CF in the fetus by antenatal detection of 2 variants, or CF in the newborn.	NA	Participant flow through the study: screen uptake, screen positive rates, diagnostic test uptake, CF disease-causing variant carriage or phenotype in the fetus/newborn. Calculation PPV and sensitivity, specificity, and NPV if comprehensive follow-up of all screened couples.	Screening pilots or cohorts. Systematic reviews of these studies.	Screening of high risk couples or cascade screening. Studies assessing analytical validity of a test to detect given variant panel. Conference abstracts, non-English language studies.
4	People in the UK: Pregnant women/couples invited for screening; people with CF carrier status; people affected by CF	CF or CF carrier status	A complete antenatal screening programme to identify a fetus with 2 disease-causing variants and so inform reproductive decisions around continuation or termination of pregnancy	NA	NA	Views on acceptability, by qualitative or quantitative assessment	Qualitative studies (interviews, focus groups etc.) with ≥10 participants, pilot studies, feasibility studies, cross-sectional or cohort studies. Systematic reviews of these studies.	Studies with <10 participants. Conference abstracts, non-English language studies.

Appraisal for quality and risk of bias

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.

Genotype-phenotype association studies for question 2 (Criterion 1) were assessed using the quality appraisal tool QUIPS. To the best of our knowledge there is no validated tool available specifically for assessing the quality of genotype-phenotype association studies. QUIPS is designed for use in prognostic studies, and is recommended by the Cochrane collaboration’s Prognosis Methods Group for assessing their risk of bias. The QUIPS tool considers 6 main components relevant to assessing risk of bias: participation, attrition, measurement of prognostic factor, outcome measurement, confounding and statistical analysis. To be consistent for all studies, the assessments were based solely on the information provided in each individual study publication. Information was not obtained from additional sources, for example through accessing national registries, study protocols or trial websites. Quality assessments for each of the individual studies included for the genotype-phenotype association question are presented in Appendix 4, Tables 30.1-15. The overall quality themes from these assessments are discussed in the discussion of findings in Criterion 1, question 2.

Diagnostic accuracy studies considered for question 3 (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. The result of this assessment is presented in Table 31 in Appendix 4, and the overall themes are discussed in the discussion of findings in Criterion 4.

Question level synthesis

Criterion 1 – Epidemiology and natural history

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 prevalence of cystic fibrosis and cystic fibrosis carrier status in the UK, overall and by mutation? How has it changed over time, and where available, what is the prevalence by ethnicity?

Background

The Murray et al HTA (1999)[3] reported an overall birth prevalence of CF of 1 in 2400 in the UK, equivalent to 300 new cases each year. Given that most CF births result from 2 parents each heterozygous for a CF-causing variant, who have a 1 in 4 chance of baby inheriting both variants, the carrier frequency is estimated at 1 in 24. A survey of 22 UK molecular laboratories reported data from analysis of 9807 chromosomes from CF carriers. The most common disease-causing variant was F508del carried by 75%, followed by G551D at 3.1%, G542X 1.6% and others with a frequency of less than 1%. Variant frequency was observed to vary across the UK.

Murray et al[3] reported that the birth prevalence and variant frequency in Asian populations was uncertain. Prevalence among those of African and Caribbean origin was reported to be very low. US studies had observed higher prevalence among black Americans than among those of black ethnicity from other countries. However this was expected to be because of the high number of black Americans with white ancestry.

This review aimed to update this information and summarise the prevalence of CF and CF carrier status in the UK, overall and by variant, and to look at how it has changed over time. Where available it also aimed to identify data by ethnicity.

Eligibility for inclusion in the review

The review aimed to look at cross sectional studies, cohort studies, surveillance reports or registry data from relevant UK populations, published since 2000. Eligible studies looked at

non-selected (for example, random or consecutively enrolled) samples of the general UK population or pregnant women, specifically. Where available the review looked at studies analysing populations by ethnicity, and looking at change over time. Systematic reviews of such studies would also be eligible.

Studies were required to include at minimum 500 people to ensure representation of the general population. Studies looking at prevalence among high-risk populations, such as people with family history of CF, were excluded. The review also excluded conference abstracts, editorials and studies not available in English language.

Description of the evidence

Twenty-two studies identified by the search were considered relevant to this question at initial appraisal, of which 11 were selected for full text appraisal. Additionally online annual reports available from the UK CF Registry were reviewed.

In May 2018 the registry had 11 annual reports available for years from 2004 to 2016, each providing data on the number of registered cases across a range of years. Three reports were selected which gave coverage of the full period: the latest available annual report from 2016 (covering 2012-16), that from 2012 (covering 2008-12) and from 2008 (covering 2002-08).[1, 2, 12] No data was available for 2000-02.

These reports provided the most up-to-date and comprehensive information on the number of registered cases, new diagnoses, and the frequency of different genotypes for each year since 2002. Any relevant studies identified by the literature search had obtained data from the UK registry and looked at the same or earlier periods. As such the primary data from the online registry source was used for this question, and 10 of the 11 studies reviewed at full text were excluded.

One study (Hoo et al[13]) was selected for inclusion as this provided data on the prevalence of cases with ‘mild’ or ‘severe’ phenotype as denoted by pancreatic sufficiency or insufficiency, respectively.

For the registry data, prevalence was calculated using mid-year population estimates from the Office for National Statistics (ONS)[14] against the number of cases registered in that year (including new diagnoses). Incidence was calculated using the new diagnoses for that year against the combined number of live births from the ONS,[15] National Records of Scotland (NRS)[16] and Northern Ireland Statistics and Research Agency (NISRA).[17]

The detailed findings extracted from the CF registry annual reports and the Hoo et al[13] study are presented in the summary and appraisal of individual studies in **Error! Reference source not found.**, Tables 23-26.

No publications reviewing carrier frequency were identified. Likewise no studies or registry data reported carrier status or CF birth frequency by ethnicity.

Discussion of findings

The UK CF Registry annual reports allowed tracking of changes in prevalence and incidence from 2002-04 and from 2007-16.[1, 2, 12] Annual reports do not cover years 2000-01 or 2005-06.

In general, there has been a steady increase in the UK prevalence of CF across the years as shown by the total cases in Figure 1, and number per 10,000 of the population in Figure 2. There were 10,461 cases registered in 2016.[1] This gives a prevalence of 1 in 6276 or 1.59 per 10,000 in 2016, compared to 1 in 8564 or 1.17 per 10,000 in 2002. The prevalence was slightly lower in 2016 compared with preceding years (it was 1.64 and 1.66 per 10,000 in 2014 and 2015, respectively) but this is said to be due to data clearing within the registry in 2016. Patients who were registered but had not had annual data submitted in that year were followed up. Those who were no longer being cared for in the NHS (given example, had moved abroad) were removed from the registry.

Looking at the number of new diagnoses, in 2002-04 there were fewer than 200 new diagnoses each year,[2] with a clear change to over 200 cases per year from 2007 onwards.[1, 12] The increase is likely due to the introduction of newborn screening. Annual reports since 2010 have documented the number of new annual diagnoses that have been identified through newborn screening, and it has accounted for between 60 to 75%. The remainder are likely identified through clinical symptoms or possibly family history, but this is not specified.

Since 2007 the number of new diagnoses each year has remained fairly stable or has not shown a clear pattern of change (most evident from Figure 2). This suggests that incidence is not increasing. Latest data from 2016 gives an incidence of CF of 1 in 3137 or 3.19 per 10,000 live births per year.[1] The overall pattern suggests a UK incidence slightly lower than general estimates of 1 in 2500 live births.

The general trend of increasing prevalence without clear increase in incidence (since screening introduction) could suggest improved survival of people with CF.

The proportion of people who have been genotyped has consistently increased since 2008. It was complete for 98% of existing cases in the registry in 2016 (excluding new diagnoses that may not have been reviewed).[1] The prevalence of CF variants has changed little over the years. The full list of all genotypes by prevalence in the 2016 annual report[1] is given in Table 12 (in relation to question 3, variants that have been included in screening pilots).

Overall it shows at least 90% of people with CF carry at least one F508del variant, with around 50% of all cases being homozygous for this variant. G551D is the next most common variant, carried by about 6% of people with CF, R117H by 4-5%, G542X by 3-4% and 621+1G→T by 2-3%. Other variants are carried by 1% or less.

Analysis of genotype frequency by UK nation shows some variation from this pattern for Scotland, Wales and Northern Ireland but this is likely due to the much smaller number of registered cases in these regions compared with England (see Appendix 3, Tables 23-25).

Data from the annual reports of the CF registry is likely to be the most reliable source of information on the prevalence and incidence of CF in the UK. The CF Trust website reports that CF care teams enter data at every specialist CF centre and clinic across the UK. Over 99% of people with CF (or their carers) are said to consent to their data being submitted. Therefore the registry is likely to include data on almost all people with CF in the UK. There is the potential for missing, inaccurate or incomplete data entry from individual centres within the UK. However, it is not possible to say from the available information how likely this may be. As mentioned above the data clearing exercise in 2016 (following up patients who had not had data submitted for that year), suggests that previous years could have had been slight overestimates. However, it is not possible to know this, and the general trend in prevalence across the years could still be similar.

There is also the possibility that prevalence figures exclude people who have inherited 2 CF variants but have mild phenotype and have not come to clinical attention. This could perhaps in part contribute to the increase in incidence since the introduction of newborn screening. There may have possibly been an increase in detection of some individuals with milder phenotype who may have had minimal symptoms and been diagnosed late in life without screening.

The final Hoo et al[13] study obtained data on the number of people in the UK CF registry in 2007-10 who were taking pancreatic enzyme replacement therapy as a sign of pancreatic insufficiency (information not contained in the annual reports). This is generally accepted to be a “severe” phenotype with pancreatic sufficiency a “mild” phenotype. Of 10,516 patients

registered during that period, the vast majority (78%) had severe phenotype by this definition and only 12% had pancreatic sufficiency (data was missing for 10%).

There are, however, some limitations to this data. Use of enzyme replacement as reported to the registry may be an imprecise indicator of pancreatic sufficiency. Evidently this information was incomplete for all patients. Additionally need for enzyme replacement may cover varying degrees of severity. The dose and duration of use may differ for individual patients. For example, it cannot inform whether the patient has been pancreatic insufficient since diagnosis in infancy or childhood, or whether this has only developed in later life. The study is also unable to inform upon the frequency of pancreatic insufficiency among new annual diagnoses. As the study covers 2007-10 this would be particularly relevant to the issue of whether newborn screening may have increased the diagnosis of milder cases. However, information on pancreatic status was missing for half of all new diagnoses so could not give a reliable indication. A final limitation is that this study gives slightly different prevalence figures for the years 2007, 08 and 09 compared to those given in the annual registry reports. The reasons for this are unclear.

Figure 1. Prevalence and incidence in total number

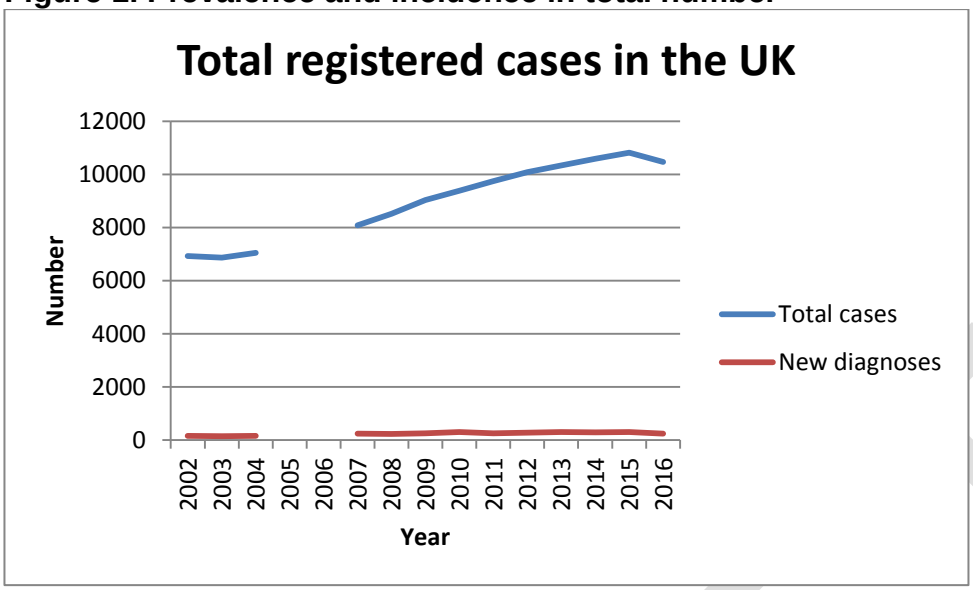
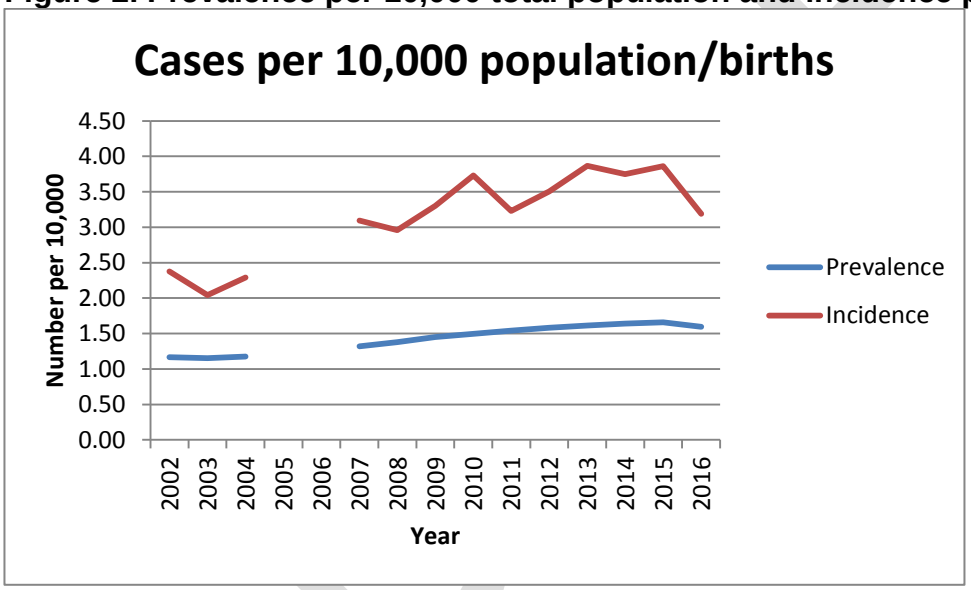


Figure 2. Prevalence per 10,000 total population and incidence per 10,000 annual births



Summary of Findings Relevant to Criterion 1 – prevalence and incidence: Criterion met‡

Data is available from annual reports from the UK CF Registry on the prevalence and incidence of CF in the UK. This shows a steady rise in prevalence since the Millennium with latest 2016 data suggesting that CF affects 1 in 6276 or 1.59 per 10,000 of the population.

There was an increase in the number of annual diagnoses coinciding with the introduction of newborn screening in 2007. However, there has since been no clear change in incidence suggesting that it is not increasing. Latest 2016 data gives a UK incidence of 1 in 3137 or 3.19 per 10,000 live births per year.

There has been little change in genotype prevalence over the years. F508del is by far the most common. Ninety percent of people with CF carry at least one copy of this variant and about half are homozygous.

99% of people with CF seen across specialist clinics across the UK are said to consent for their data to be submitted to the UK CF registry. Therefore this information is likely to represent all people with CF in the UK, barring any potential errors from missing, inaccurate or incomplete data entry. Therefore, this part of criterion 1 on the prevalence and incidence of CF in the UK is met.

There was no data on prevalence or incidence by ethnicity in the UK, nor any recent data on carrier prevalence in the UK.

‡ ‡**Met** -for example, this should be applied in circumstances in which there is a sufficient volume of evidence of sufficient quality to judge an outcome or effect which is unlikely to be changed by further research or systematic review.
Not Met - for example, this should be applied in circumstances where there is insufficient evidence to clearly judge an outcome or effect or where there is sufficient evidence of poor performance.
Uncertain -for example, this should be applied in circumstances in which the constraints of an evidence summary prevent a reliable answer to the question. An example of this may be when the need for a systematic review and meta-analysis is identified by the rapid review.

Question 2 – What are the genotype-phenotype associations in cystic fibrosis patients, including their clinical prognosis?

Background

The causative *CFTR* gene was first described in 1989.[18] At that time it was recognised that people with CF fell into two groups: those who are pancreatic sufficient and pancreatic insufficient, the latter of whom formed the largest clinical subgroup. The researchers also observed that people with insufficiency tended to be far more homogenous in terms of variants than those with sufficiency who were more heterogeneous.[18] The most common disease-causing variant was identified as F508del, which causes a deletion of phenylalanine at amino acid position 508 of the protein. As demonstrated in question 1, in the UK about 90% of people with CF carry at least 1 copy of this variant and about 50% carry 2 copies.

Further research during the 1990s[19-22] built on these observations and categorised variants into 5 classes depending on their position in the *CFTR* gene and the functional effect that this had on the protein product.[6, 10]

1. Protein production variants. These are typically nonsense, frameshift, large deletions/insertions or splice variants that cause absent or abnormal CFTR protein production. The resulting effect is that no functional protein is made and therefore none is present in the membrane.
2. Protein processing variants. These are variants that affect how CFTR is processed within the cell and transported to the membrane. The variants tend to cause amino acid deletions or substitutions which stop the CFTR protein from folding correctly. Although potentially functional protein may be produced there is no protein present in the membrane. F508del typifies this class.
3. Regulation/gating variants. These variants prevent ATP binding and hydrolysis at the nucleotide-binding domains of the CFTR protein, which is required for channel activity. Although a normal amount of protein is present in the membrane it is non-functioning. An example is the common variant G551D. A treatment (ivacaftor) is now available for variants of this class which increases activity of the ion transport channel.
4. Conduction variants. These variants reduce the ability of chloride to flow through the channel but there is still some residual function. These variants typically occur in the region of the gene that encodes the first membrane-spanning domain of the protein.
5. Reduced production or processing variants. These variants cause reduced production or trafficking of CFTR within the cell. The resulting effect is that functioning protein is present in the membrane but in reduced quantity.

As class1 to 3 variants result in absent or non-functioning protein, individuals carrying 2 of these variants would be expected to have a severe disease course. This is likely to be the case for the large number of people with CF who will be homozygous for F508del. The presence of at least one class 4 or 5 variant may result in enough functioning protein in the

membrane to confer a milder disease course, even if it is present alongside a class 1-3 variant.[10]

However, the Murray et al HTA[3] concluded that it can be very difficult to predict the likely disease course even for people with 2 severe class 1-3 variants. The clinical phenotype can vary widely in people with CF with inconsistent effects across genotypes. Many other variables may also have an influence on disease course complicating phenotype predictions, such as complex alleles, modifier genes (other than *CFTR*); environmental factors; and care and treatment availability, particularly the development of disease-specific treatment.

In the context of antenatal screening, a clear understanding of whether a particular genotype can reliably predict the expected clinical course of disease would be essential as it would be used to guide informed decision making. Therefore a systematic review was conducted which aimed to explore whether a consistently predictive association between *CFTR* genotype and phenotype can be established. Where possible the review also aimed to see whether there is any evidence that neonatal screening or treatment advances have altered clinical prognosis.

Eligibility for inclusion in the review

The broad inclusion criteria for this question (outlined in Table 2) was decided *a priori* in discussion with UK NSC and an expert CF advisor.

The reviewers included studies that performed an analysis to look at the relationship between *CFTR* genotype and phenotype. This could include prospective or retrospective cohort or cross-sectional studies comparing clinical outcomes in people with CF of different *CFTR* genotype. It could also include case-control studies comparing *CFTR* genotypes in people with and without a specific clinical outcome (for example, pancreatic insufficiency). Systematic reviews of such studies would also be eligible.

No restrictions were placed on study date, country, screening setting, or the type of care that could be provided to patients. The reviewers reported and considered the impact of these factors when identified in eligible studies.

The clinical outcomes to be considered were decided in discussion with the topic advisor, in order to focus upon outcomes likely to be of greatest importance to people with CF, and therefore be of most relevance to parents making reproductive decisions.

With this in mind, the primary outcomes of interest were:

- Survival, life expectancy
- Age at diagnosis (as an indication of symptom severity)
- Respiratory function and infection
- Pancreatic sufficiency and nutritional status
- Treatment burden (for example, the number of treatments received)
- Quality of life (for example, time off school or repeated hospitalisation)

Studies of solely biochemical/physiological outcomes such as sweat chloride levels were not included, as they are unlikely to be sufficient to inform reproductive decision making.

The review has only looked at the relationship between *CFTR* genotype and clinical phenotype. While other variables may affect phenotype, these were outside of the scope of this review.

The review did not include studies looking at:

- the role of genes other than CFTR
- the role of environmental modifiers
- how cellular or molecular factors (for example, immune cells or cytokines) may modulate disease outcomes
- the modifying effect of colonisation with infectious organisms (such as MRSA) upon disease course
- carriage of *CFTR* variants in people with “atypical” or “non-classic” CF which tends to include single-organ manifestations or CFTR-related conditions (such as male infertility/absent vas deferens, pancreatitis or respiratory conditions)
- the effect of complex alleles (more than one disease-causing variant on the same allele)
- rare variants carried by <0.1% of the UK population with CF (as guided by the latest report from the UK CF registry[1] – see Table 12)

These latter 2 exclusions were based on the rationale that potential antenatal screening programmes would be likely to screen for a selected panel of the more common variants rather than sequence the full *CFTR* gene. As covered by question 3 of this report, all screening pilots identified to date have taken this approach.

The following studies were also excluded from the review:

- Individual case reports or case series
- Cohorts with an initial study sample size of fewer than 100 people (on the premise that smaller studies may be less reliable in identifying genotype-phenotype associations)
- Studies reporting variant frequency in people with CF but not reporting any assessment of link with phenotype

- Studies reporting clinical outcomes/phenotype of people with CF but not reporting any assessment of link with genotype
- Studies not looking at the clinical outcomes of interest, including those looking at differences by genotype at the cellular/molecular level (including channel activity) or differences in sweat chloride
- Treatment trials reporting drug response in people with particular genotype but not comparing response in people with other genotype

The latter exclusion was based on the rationale that the question aimed to address whether certain genotypes may need more/less treatment or respond more/less favourably to treatment. However, it was not looking at whether there is an effective treatment available for a specific genotype.

Finally the reviewers excluded studies not available in English language, editorials or non-systematic literature reviews, conference abstracts, or letters.

Description of the evidence

Searches were performed in MEDLINE and EMBASE databases (Embase.com), The Cochrane Library (Wiley Online) and Scopus on 11th May 2018, with no date restriction. The full search strategy is presented in 0 alongside a PRISMA diagram which outlines the flow of studies at each stage of appraisal.

Database searches yielded 9238 results, which were filtered at title and abstract level by an information specialist. Of these, 841 were considered potentially relevant to this question and were further reviewed at abstract by the main analyst. Applying the exclusion criteria as described above, 76 studies were selected for full text appraisal. These studies were reviewed by the main reviewer and potential inclusions and exclusions were discussed with a second senior reviewer. The final study selections were checked for any potential omissions by the topic expert advisor. Any additional studies suggested at this stage were checked against the review's inclusion criteria, and added if they met these. Citations of included studies were also hand-searched, although this process identified studies of contextual relevance only.

Of the studies accessed at full text, 47 studies were excluded. Reasons for exclusion were varied. This included studies that did not examine the clinical outcomes of interest, or simply gave a long list of the individual genotypes of people with a specific clinical outcome which prevented meaningful analysis. A full list of the excluded studies with their individual reason for exclusion is provided in Appendix 2, Table 22.

The remaining 29 studies met inclusion criteria.

Fifteen studies were selected to provide the main evidence examining the link between genotype and phenotype, and were extracted in detail. These studies mostly analysed data from national registries or international consortiums, thereby including several thousand people. In a few of these studies the sample sizes for analysis became smaller when identifying people within the registry who had specific genotype. However, because the design of these studies meant that they would be expected to represent all people with these genotypes from the assessed country or region (in the case of European or international consortiums) they were prioritised for inclusion. The findings from these studies are summarised in Tables 3-7 according to the clinical outcomes assessed and indicating the genotype comparison performed. A full extraction of data from each of these individual studies is presented in Appendix 3, Table 27. Quality appraisal is presented in Appendix 4 and summarised in Table 8.

The findings of an additional 14 non-prioritised studies are summarised in Appendix 3, Table 28. These are studies from single centres or a few regional centres. The majority of these studies are too small to provide reliable statistical analysis but have been included to assess whether the broad direction of effect is consistent with the registry studies.

The studies fell into 3 main groups of genotype comparison:

- by class of variant “severe” (both variants class 1 to 3) vs “mild” (≥1 variant class 4 or 5)
- F508del homozygotes vs F508del heterozygotes or non-F508del heterozygotes
- by comparison of specific genotypes

All of these comparisons could provide information of relevance to potential antenatal screening programmes.

- All variants in the ACMG antenatal screening panel and the majority of those included in other screening pilots to date have now been classified 1 to 5 (see question 3, Table 12).
- The vast majority of people with CF will be either homozygous or heterozygous for F508del
- Specific variants assessed by the identified studies have almost all been included in antenatal screening panels (with the exception of one study analysing an unclassified variant, P67L[23]).

No studies assessing quality of life in people with CF and different genotype were identified. Only a single study compared treatment burden in people with different genotype.[24] No other studies assessed or reported treatments that had been given to the included

individuals (aside from enzyme replacement therapy in the context of assessments of pancreatic sufficiency).

Due to the heterogeneity of studies in terms of their methods of genotype comparison, outcome assessment and analysis, the decision was made not to perform meta-analysis. The findings of the studies have been discussed and summarised narratively to show the range of results for different outcomes and the overall direction of effect. Similarly the decision was made not to perform further statistical analysis when the study authors provided only the comparative proportions of people with different clinical outcome according to variant class or by F508del homozygosity or heterozygosity. Within the confines of the study publication and the associated quality limitations, it was not thought appropriate to calculate predictive risk ratios that may not be reliable and informative when the authors accessing the primary data had not considered it appropriate to do so.

Table 3. Relationship between genotype and survival

Study	Design and Setting	Population	Genetic comparison	Findings
McKone et al 2006[25]	Retrospective cohort US CF Foundation Registry 1993 to 2002	N=15,651	Severe vs mild class (both variants class 1-3 vs ≥1 in class 4-5)	Increased mortality for severe vs mild class variants Median survival 36 years for severe vs 50 years mild Median age at death (for n=1672 who died): 24.2 vs 37.6 years Adjusted hazard ratio (aHR) for mortality: severe genotype 1.60 (95% confidence interval [CI] 1.20 to 2.10) Adjustment for year of entry to the cohort, population size of the CF centre, age, and phenotypic variables Severe genotype as predictor of death <30 years Sensitivity 98%, Specificity 11%, PPV 69%, NPV 71%
McKone et al 2003[26]	Retrospective cohort US CF Foundation Registry 1991 to 1999	N=17,853	F508del/F508del vs F508del/other variant (11 most common in registry) Class 2/2 (mostly F508del/F508del) vs class 2/other class	Certain F508del heterozygotes have reduced mortality vs homozygotes F508del/F508del standardised mortality ratio (SMR) 21.8 (95% 20.5 to 23.1) (adjusted for age and gender) Genotypes with lower SMRs than F508del/F508del (p<0.01): F508del/I507del SMR 8.0 (95% 2.7 to 13.3) F508del/R117H SMR 4.7 (95% CI 0.8 to 8.5) F508del/3849+10kbC>T SMR 11.9 (95% CI 5.0 to 18.9) F508del/2789+5G>A SMR 4.4 (0.0 to 8.9) No significant difference for F508del heterozygotes with: G551D, G542X, N1303K, W1282X, R553X, 621+1G>T and 1717+1G>A One variant class 4 or 5 gives reduced mortality vs both variants class

Study	Design and Setting	Population	Genetic comparison	Findings
				2 (p<0.0001) Class 2/2 SMR 21.2 (95% CI 20.0 to 22.5) Class 4 SMR 7.8 (95% CI 4.2 to 11.4) Class 5 SMR 9.1 (95% CI 4.8 to 13.5) No significant difference class 1 or 3
Lai et al 2004[27]	Retrospective cohort US CF Foundation Registry 1986 to 2000	N=13,690	F508del/F508del vs 2 severe class variants (including F508del/other) vs ≥1 mild class variant	Reduced risk of "shortened" survival for mild class and other severe class vs F508del homozygotes Severe genotype: odds ratio (OR) 0.76 (95% CI 0.67 to 0.86) Mild genotype: OR 0.51 (95% CI 0.37 to 0.70)
O'Connor et al 2002[28]	Retrospective cohort US CF care centres 1982 to 1998	N=15,214	F508del/F508del vs F508del/other vs other/other	Increased mortality risk for F508del homozygotes and those with 2 non-F508del variants vs F508del heterozygotes F508del/F508del: aHR 1.36 (95% CI 1.19 to 1.55) Other/other: aHR 1.40 (95% CI 1.15 to 1.71) Adjusted for gender, age and type of presentation, ethnicity and socioeconomic status
Simmonds et al 2009[29]	Case control UK single centre patients born 1965 surviving to 2004 vs UK adult CF registry	N=112 >40 years N=3989 adults (aged >16) in the registry	F508del/F508del vs F508del/other	F508del homozygotes are less common among older patients F508del/F508del: 30% patients aged >40 years vs 50% of the total registry population (p<0.001) F508del heterozygotes with an unknown second variant are more common among older patients F508del/"unknown"*: 32% patients aged >40 years vs 13% of the total

Study	Design and Setting	Population	Genetic comparison	Findings
	population			registry population (p<0.001) *No significant difference between older patients and the total registry population for F508del/"known" variants of R117H, R347P, G551D, G542X, N1303K, G85E, 1717-1G>A and 621+1G>T (pooled as a collective group: 14 vs 22%, p=0.06)
Badet et al 2004[30]	Case control French registry patients born <1970 and >30 years in 2000 vs remaining French registry population	N=114 >30 years N=3220 registry population	F508del/F508del vs F508del/other vs other/other	No difference in genotype between patients aged >30 years and the wider patient registry population F508del/F508del: 56% patients aged >30 vs 58% registry population F508del/other: 33% vs 21% other/other: 11% vs 21% (P>0.05)

Table 4. Relationship between genotype and pancreatic sufficiency and nutritional status

Study	Design and Setting	Population	Genetic comparison	Findings
McKone et al 2003[26]	Retrospective cohort US CF Foundation Registry 1991 to 1999	N=17,853	F508del/F508del vs F508del/specific variant* Class 2/2 (mostly F508del/F508del) vs class 2/other class *22 variants, 21 of which are compatible with ACMG 2004 panel	Certain F508del heterozygotes have lower rates of pancreatic insufficiency and greater weight than homozygotes All below values are described to be means as expected for a 15 year old in a cohort where 52% were male F508del/F508del pancreatic insufficiency 92% (95% CI 91-92), height 141cm (+/- 0.2) and weight 37.0kg (+/- 0.1) Genotypes with improved status:

Study	Design and Setting	Population	Genetic comparison	Findings
			with the exception of S549N (assessed here and not covered in panel) and 3120+1G>A (added to the panel and not covered here)	F508del/I507del pancreatic insufficiency 84% (78-89) F508del/R117H insufficiency 65% (55-73), weight 42.9kg (+/- 1.7) F508del/3849+10kB insufficiency 66% (57-74), weight 41.2kg (+/- 1.2) F508del/2789+5G>A insufficiency 71% (59-81) F508del/R347P insufficiency 67% (52-79) F508del/A455E insufficiency 60% (41-76) F508del/R334W insufficiency 67% (46-82) All remaining variants no difference in risk (all p<0.001) <u>Lower rates of pancreatic insufficiency and greater body weight for one variant class 4 or 5 vs both variants class 2</u> Class 2/2: insufficiency 92% (91-93), height 141cm (+/- 0.2), weight 37.0kg (+/- 0.1) Class 4: insufficiency 71% (64-76), weight 41.0kg (+/- 1.1) Class 5: insufficiency 68% (61-74), weight 41.5kg (+/- 1.0) (all p<0.001) <u>NB also improved status for unidentified and unclassified variants compared with F508del homozygotes</u> “Unclassified” variant: insufficiency 84% (83-85), weight

Study	Design and Setting	Population	Genetic comparison	Findings
				38.2kg (+/- 0.2) F508del/other unidentified variant: insufficiency 86% (84-87), weight 38.1kg (+/- 0.3) other unidentified/other unidentified variant: insufficiency 81% (80-84), weight 38.3kg (+/- 0.3)
Koch et al 2001[31]	Cross sectional European Epidemiologic Registry of CF (ERCF), 9 countries 1997: patients with 180 days follow-up but first assessment of variable taken	N=8963	Comparison across class of both variants	Patients with at least one class 4 variant have lower rates of pancreatic insufficiency and higher weight-for-age percentiles Mean values are given; no statistical analysis performed Pancreatic insufficiency Class 4/any other variant Under 18s: 71.3% Over 18s: 52.3% Class 1/1, 2/2, 2/3 (ranges across these 3 groups): Under 18s: range 96.4 to 98.0% Over 18s: range 95.8 to 100% Weight-for-age percentile Class 4/any variant: Under 18s: 42.3 Over 18s: 44.3 Class 1/1, 2/2, 2/3 (range across groups): Under 18s: range 25.9 to 39.0 Over 18s: range 14.0 to 26.8

Study	Design and Setting	Population	Genetic comparison	Findings
				NB: assessed for patients available in the registry. Few patients had class 3/3 or 5/any for reliable comparison; no analysis was available for those with class1/2 variants
Dewulf et al 2015[24]	Retrospective cohort Belgian CF Registry 2010	N=747	Severe vs mild class (both variants class 1-3 vs ≥1 in class 4-5)	Severe class variants have higher rates of pancreatic insufficiency than mild class Pancreatic insufficient: severe 98.8% vs mild 36.5%, p<0.001
Green et al 2010[32]	Retrospective cohort US CF Twin and Sibling Study (CFTSS) Followed after enrolment (date unclear) to Dec 2008	N=1381	Severe vs mild class (both variants class 1-3 vs ≥1 in class 4-5)	Severe class variants have higher rates of pancreatic insufficiency than mild class Pancreatic insufficient: severe 97.8% vs mild 30.3%, p<0.001
Radtke et al 2017[33]	Cross sectional International, multicentre members of the European CF Society 32 centres	N=726	Severe vs mild class (both variants class 1-3 vs ≥1 in class 4-5)	Severe class variants have higher rates of pancreatic insufficiency, lower BMI and lower %body fat than mild class Pancreatic insufficiency: severe 95% vs mild 24%, p<0.05 BMI z score: severe -0.25 (95% CI -0.95 to +0.42) vs mild -0.11 (95% CI -0.77 to +0.74), p<0.05 (Number of standard deviations below the mean for age and sex) Body fat: severe mean 18.2% (+/- 5.7) vs mild mean 21.8% (+/- 6.4), p<0.001

Study	Design and Setting	Population	Genetic comparison	Findings
The Cystic Fibrosis Genotype-Phenotype Consortium 1993[34]	Cross sectional 32/89 centres belonging to the CF Genetic Analysis Consortium Time period unclear	N=399 F508del/F508del N=399 F508del/other Age- and sex- matched from the same centre	F508del/F508del vs F508del/specific variant: G542X, R553X, W1282X, N1303K, R117H, 621+1G>T, 1717-1G>A	No difference in rates of pancreatic insufficiency between F508del homozygotes and any compound heterozygotes with exception of R117H F508del/F508del insufficiency 96% vs F508del/R117H 13%, p<0.001
Dugueperoux and De Braekeleer 2005[35]	Cross sectional French CF registry patients who attended participating centres 1992 to 2002 and carrying variants 3849+10kbC>T or 2789+5G>A	N=16 F508del/F508del N=16 F508del/3849+10kbC>T N=34 F508del/F508del N=34 F508del/2789+5G>A Age- and sex-matched from the same centre	Specific genotype comparison: F508del/F508del vs F508del/3849+10kbC>T vs F508del/2789+5G>A	Lower rates of pancreatic insufficiency among 2789+5G>A and 3849+10kbC>T heterozygotes compared with F508del homozygotes F508del/F508del insufficiency 100% F508del/3849+10kbC>T 46.6%, p=0.002 F508del/F508del insufficiency 97.0% F508del/2789+5G>A insufficiency 59.4%, p=0.002
MacKenzie et al 2017[23]	Retrospective cohort Canadian CF registry patients who attended clinics 1996 to 2011 and carrying the P67L variant	N=266 F508del/F508del N=26 F508del/P67L	Specific genotype comparison: F508del/F508del vs F508del/P67L	Lower rates of pancreatic insufficiency among P67L heterozygotes compared with F508del homozygotes F508del/F508del insufficiency 99% F508del/P67L insufficiency 26.9%, p<0.001

Table 5. Relationship between genotype and lung function and infection

Study	Design and Setting	Population	Genetic comparison	Findings
McKone et al 2003[26]	Retrospective cohort US CF Foundation Registry 1991 to 1999	N=17,853	F508del/F508del vs F508del/specific variant* Class 2/2 (mostly F508del/F508del) vs class 2/other class *22 variants, 21 of which are compatible with ACMG 2004 panel with the exception of S549N (assessed here and not covered in panel) and 3120+1G>A (added to the panel and not covered here)	<u>Certain F508del heterozygotes have improved lung function and less *P. aeruginosa* colonisation compared with homozygotes</u> (FEV1 is forced expiratory volume in 1 second. FVC is forced vital capacity. Values are means) F508del/F508del FEV1 77% predicted (+/- 0.3), FVC 89% predicted (+/- 0.3), *P. aeruginosa* colonisation 60% (95% CI 59-61) Genotypes with improved lung function and less infection: F508del/I507del FEV1 86% (+/- 2.1), *P. aeruginosa* 39% (31-48) F508del/R117H FEV1 91% (+/- 2.1), FVC 97% (+/- 1.7), *P. aeruginosa* 22% (16-29) F508del/2789+5G>A FEV1 88% (+/- 2.8), FVC 97% (+/- 2.3), *P. aeruginosa* colonisation 32% (22-44) F508del/560T FEV1 84% (+/- 3.3) F508del/A455E FEV1 98% (+/- 4.0), FVC 104% (+/- 3.4), *P. aeruginosa* 17% (8-32) All remaining variants no difference in risk (all p<0.001) <u>Improved lung function and less *P. aeruginosa* colonisation for one variant class 4 vs both variants class 2</u> Class 2/2 FEV1 78% predicted (+/- 0.3), FVC 89% predicted (+/- 0.3), *P. aeruginosa* colonisation 59% (58-

Study	Design and Setting	Population	Genetic comparison	Findings
				61) Class 4 FEV1 85% (+/- 1.4), FVC 94% (+/- 1.2), *P. aeruginosa* 37% (31-43) (all p<0.001) <u>NB also improved lung function and reduced infection for unidentified and unclassified variants compared with F508del homozygotes</u> "Unclassified" variant: FEV1 81% (+/- 0.4), FVC 90% (+/- 0.4), *P. aeruginosa* 46% (44-48) F508del/other unidentified variant: FEV1% 80 (+/- 0.5), FVC 91% (+/- 0.5), *P. aeruginosa* 50 (48-52) other unidentified variant/other unidentified variant: FEV1 82% (+/- 0.6), *P. aeruginosa* 40% (38-43)
Lai et al 2004[27]	Retrospective cohort US CF Foundation Registry 1986 to 2000	N=3320 for lung function N=5290 for *P. aeruginosa* colonisation	F508del/F508del vs 2 severe class variants (including F508del/other) vs ≥1 mild class	<u>No difference in lung function or infection for severe and mild variants vs F508del homozygotes</u> FEV1<70% Severe genotype: OR 0.88, 95% CI 0.74 to 1.05 Mild genotype: OR 1.16, 95% CI 0.55 to 1.33 *P. aeruginosa* colonisation: Severe genotype: OR 1.03, 95% CI 0.95 to 1.11 Mild genotype: OR 0.65, 95% CI 0.42 to 1.00
Koch et al 2001[31]	Cross sectional European	N=8963	Comparison across class of both variants	<u>Patients with at least one class 4 variant have slightly better FEV1 and less *P. aeruginosa* colonisation</u>

Study	Design and Setting	Population	Genetic comparison	Findings
	Epidemiologic Registry of CF (ERCF), 9 countries 1997: patients with 180 days follow-up but first assessment of variable taken			Mean values are given, statistical analysis was not performed FEV1 % predicted Class 4/any variant: Under 18s: 82.8% Over 18s: 61.8% Class 1/1, 2/2, 2/3: Under 18s: range 71.3 to 78.9% Over 18s: range 50.2 to 58.0% *P. aeruginosa* % colonisation Class 4/any variant: Under 18s: 33% Over 18s: 56.7% Class 1/1, 2/2, 2/3: Under 18s: range 50.0 to 55.1% Over 18s: range 81.7 to 100% Minimal difference in FVC Class 4/any variant: Under 18s: 89.4% Over 18s: 76.5% Class 1/1, 2/2, 2/3: Under 18s: range 85.5 to 88.3% Over 18s: range 67.4 to 74.1%

Study	Design and Setting	Population	Genetic comparison	Findings
				NB: assessed for patients available in the registry. Few patients had class 3/3 or 5/any for reliable comparison; no analysis was available for those with class1/2 variants
Dewulf et al 2015[24]	Retrospective cohort Belgian CF Registry 2010	N=747	Severe vs mild class (both variants class 1-3 vs ≥1 in class 4-5)	Severe class poorer lung function and greater *P. aeruginosa* colonisation than mild class FEV1 % predicted: severe 77.0% (IQR 55.6 to 94.1) vs mild 86.8% (IQR 68.1 to 103.0), p<0.001 Chronic *P. aeruginosa* infection: severe 36.2% vs mild 14.1%, p<0.001
Green et al 2010[32]	Retrospective cohort US CF Twin and Sibling Study (CFTSS) Followed after enrolment (date unclear) to Dec 2008	N=1381	Severe vs mild class (both variants class 1-3 vs ≥1 in class 4-5)	Severe class have higher risk of *P. aeruginosa* colonisation than mild class using any definition First infection (+ve culture, prior -ve): HR 3.17 (95% CI 2.10 to 4.78) Chronic infection (3 +ve cultures in 6 months): HR 5.47 (95% CI 2.20 to 13.58) Multiple infections (3+ve without time definition): HR 3.81 (95% CI 2.32 to 6.28) Persistent infection (+ve cultures in ≥2 of 3 consecutive years): HR 3.32 (95% CI 2.00 to 5.50) Adjusted for FEV1 and number of cultures performed (ethnicity and gender had been assessed but were not adjusted as they did not have significant effect on lung infection) Also poorer lung function FEV1: severe 0.68 (±0.26) vs 0.75 (±0.25), p<0.001
Radtke et al	Cross sectional	N=726	Severe vs mild class	Greater *P. aeruginosa* colonisation for severe than mild

Study	Design and Setting	Population	Genetic comparison	Findings
2017[33]	International, multicentre members of the European CF Society 32 centres		(both variants class 1-3 vs ≥1 in class 4-5)	class but no difference in lung function FEV1 % predicted: severe 79 (95% CI 59 to 93) vs mild 84 (95% CI 68 to 96) (ns) *P. aeruginosa* infection %: severe 54 vs mild 36, p<0.001 Also no difference in main study outcomes of peak oxygen uptake and maximum work rate
The Cystic Fibrosis Genotype-Phenotype Consortium 1993[34]	Cross sectional 32/89 centres belonging to the CF Genetic Analysis Consortium Time period unclear	N=399 F508del/F508del N=399 F508del/other Age- and sex- matched from the same centre	F508del/F508del vs F508del/specific variant: G542X, R553X, W1282X, N1303K, R117H, 621+1G>T, 1717-1G>A	No significant difference in FEV1 between F508del homozygotes and any compound heterozygotes
Szczesniak et al 2017[36]	Retrospective cohort US CF Foundation Patient Registry Patients with repeat FEV1 data recorded 1997 to 2013	N=18,387	F508del/F508del vs F508del/other vs other/other	Patients not carrying F508del have increased risk of early sustained FEV1 decline vs homozygotes Other/other: OR 1.73 (95% CI 1.36 to 2.21) Model with adjustment for gender, age at diagnosis, birth cohort year, socioeconomic status and phenotypic variables
De Boeck and Zolin 2017[37]	Retrospective cohort European CF	N=11,417	F508del/F508del vs variants of: Class 1 and class 1/2	Having one variant class 4 or 5 confers gives less annual FEV1 decline than other groups Mean annual decline FEV1 % predicted:

Study	Design and Setting	Population	Genetic comparison	Findings
	Society Patient Registry (ECFSPR), 12 countries Repeat assessments collected 2008, 09 and 10		Class 3 and class 1/2/3 Class 4 and class 1/2/4 Class 5 and class 1/2/5	F508del/F508del -1.52% (-1.72 to -1.31) at least one class 1 variant -1.35% (-1.70 to -0.99) at least one class 3 variant -1.24% (-1.87 to -0.61) at least one class 4 variant -0.62% (-1.30 to +0.06) at least one class 5 variant -0.35% (-1.21 to +1.0) Pooled groups of those with at least one class 4 or 5 variant have small difference of +0.88% in yearly change compared to the other three groups (p=0.004) Analysis restricted to those with baseline FEV1>90% showed greatest difference for F508del homozygotes, class 1 and 3 (range -4.00 to -4.28) vs class 4 or 5 (-1.78 to -1.88)
Dugueperoux and De Braekeleer 2005[35]	Cross sectional French CF registry patients who attended participating centres 1992 to 2002 and carrying variants 3849+10kbC>T or 2789+5G>A	N=16 F508del/F508del N=16 F508del/3849+10kbC>T N=34 F508del/F508del N=34 F508del/2789+5G>A Age- and sex-matched from the same centre	Specific genotype comparison: F508del/F508del vs F508del/3849+10kbC>T vs F508del/2789+5G>A	F508del/2789+5G>A better FEV1 vs F508del homozygotes F508del/F508del FEV1 59.06% (+/- 24.87) F508del/2789+5G>A FEV1 75.38 (+/- 29.69), p=0.03 No difference for FVC No difference in lung function for F508del/3849+10kbC>T vs F508del homozygotes

Table 6. Relationship between genotype and age at diagnosis

Study	Design and Setting	Population	Genetic comparison	Findings
McKone et al 2003[26]	Retrospective cohort US CF Foundation Registry 1991 to 1999 *Screening context not reported*	N=17,853	F508del/F508del vs F508del/specific variant* Class 2/2 (mostly F508del/F508del) vs class 2/other class *22 variants, 21 of which are compatible with ACMG 2004 panel with the exception of S549N (assessed here and not covered in panel) and 3120+1G>A (added to the panel and not covered here)	<u>Certain F508del heterozygotes are diagnosed at an older age than homozygotes</u> (mean values) F508del/F508del age at diagnosis 2.5 years (+/-0.1) Genotypes associated with a significantly later diagnosis: F508del/G551D 3.7 years (+/- 0.3) F508del/I507del 8.5 years (+/- 1.1) F508del/R117H 13.7 years (+/- 1.2) F508del/3849+10kbC>T 11.3 years (+/- 0.9) F508del/2789+5G>A 13.4 years (+/- 1.6) F508del/G85E 9.2 years (+/- 1.8) F508del/A455E 14.3 years (+/- 2.0) F508del/R334W 13.2 years (+/- 3.0) (all p<0.001) No difference in risk for all remaining F508del heterozygotes. <u>Earlier age at diagnosis with one variant class 1 and increased age for one variant class 4 or 5 vs both variants class 2</u> Class 2/2 age at diagnosis 2.6 (+/- 0.1) Class 1 age at diagnosis 2.0 (+/- 0.1)

Study	Design and Setting	Population	Genetic comparison	Findings
				Class 4 age at diagnosis 11.4 (+/- 0.8) Class 5 age at diagnosis 12.6 (+/- 0.7) (all p<0.001) <u>NB also increased age at diagnosis for unidentified and unclassified variants compared with F508del homozygotes</u> "Unclassified" variants: age at diagnosis 6.4 (+/- 0.1) F508del/other unidentified variant: age at diagnosis 5.8 (+/- 0.2) Other unidentified/other unidentified variant: age at diagnosis 7.5 (+/- 0.3)
Dewulf et al 2015[24]	Retrospective cohort Belgian CF Registry Patients enrolled 2010 *No screening*	N=747	Severe vs mild class (both variants class 1-3 vs ≥1 in class 4-5)	<u>Earlier age at diagnosis for patients with mild class variants</u> Severe 0.3 years (interquartile range [IQR] 0.1 to 1.3) vs mild 5.2 years (IQR 0.4 to 20.9), p<0.001
The Cystic Fibrosis Genotype-Phenotype Consortium 1993[34]	Cross sectional 32/89 centres belonging to the CF Genetic Analysis Consortium Time period unclear	N=399 F508del/F508del N=399 F508del/other Age- and sex- matched from the same centre	F508del/F508del vs F508del/specific variant: G542X, R553X, W1282X, N1303K, R117H, 621+1G>T, 1717-1G>A	<u>No difference in age at diagnosis between F508del homozygotes and F508del heterozygotes with exception of F508del/R117H</u> F508del/F508del mean age at diagnosis 2.5 (+/- 4.3) vs F508del/R117H mean 10.2 years (+/- 10.5), p=0.002 (R117H was the only mild class 4/5 variant assessed)

Study	Design and Setting	Population	Genetic comparison	Findings
	Screening context not reported			
Dugueperoux and De Braekeleer 2005[35]	Cross sectional French CF registry patients who attended participating centres 1992 to 2002 and carrying variants 3849+10kbC>T or 2789+5G>A *No screening*	N=16 F508del/F508del N=16 F508del/3849+10kbC>T N=34 F508del/F508del N=34 F508del/2789+5G>A Age- and sex-matched from the same centre	Specific genotype comparison: F508del/F508del vs F508del/3849+10kbC>T vs F508del/2789+5G>A	<u>F508del homozygotes diagnosed at earlier age than both mild class F508del heterozygotes assessed</u> F508del/F508del mean age at diagnosis 3.1 years (+/- 5.1) F508del/3849+10kbC>T mean 12.7 years (+/- 9.6), p=0.002 F508del/2789+5G>A mean 16.6 years (+/- 12.7), p=0.0001
MacKenzie et al 2017[23]	Retrospective cohort Canadian CF registry patients who attended clinics 1996 to 2011 and carrying the P67L variant *No screening*	N=266 F508del/F508del N=26 F508del/P67L	Specific genotype comparison: F508del/F508del vs F508del/P67L	<u>F508del homozygotes diagnosed at an earlier age than P67L heterozygotes</u> F508del/F508del mean age at diagnosis 0.92 years (+/- 0.13) F508del/P67L mean 18.2 years (+/- 14.6), p<0.001

Table 7. Relationship between genotype and treatment burden

Study	Design and Setting	Population	Genetic comparison	Findings
Dewulf et al 2015[24]	Retrospective cohort Belgian CF Registry 2010	N=747	Severe vs mild class (both variants class 1-3 vs ≥1 in class 4-5)	Mild class need fewer therapies over the course of one year than severe class Assessed by treatment burden index (TBI) - weighting of number of low, medium and high intensity therapies) Median TBI: severe 9 (IQR 6-12) vs mild 6 (IQR 3-8) Significant effect of variant class in regression analysis adjusted for age, gender and FEV1 (p<0.001): Mild class 23.1% decrease in TBI (95% CI 15.0 to 30.5) compared with severe class Proportion hospitalised: 50.8% severe vs 24.7% mild, p<0.001 Use of IV antibiotics: 46.0% severe vs 23.5% mild, p<0.001

Discussion of findings

Overall assessment of quality and applicability

The summary risk of bias for each of the domains of QUIPS is displayed in Table 8 and a summary of the overall quality themes across the studies is presented below. Full assessments for each individual study are presented in Appendix 4, Table 30.1-15 and a more detailed discussion of the QUIPS quality assessments by domain with accompanying rationale is presented in Appendix 5. The few quality issues that were specific to individual outcomes are presented in the following section along with the findings by outcome. However, as most studies analysed multiple outcomes, the quality issues generally apply across outcomes with little difference by outcome.

Table 8. QUIPS assessments for genotype-phenotype association studies

Study	**Summary risk of bias by domain**					
	Participation	Attrition	Genotype measure	Phenotype measure	Confounding	Statistical analysis
McKone et al 2006[25]	moderate	high	moderate	moderate	moderate	low
McKone et al 2003[26]	low	moderate	moderate	moderate	high	low
Lai et al 2004[27]	moderate	high	high	High	high	moderate
O'Connor et al 2002[28]	moderate	high	high	High	moderate	low
Simmonds et al 2009[29]	high	high	high	moderate	high	high
Badet et al 2004[30]	high	moderate	high	Low	high	high
Koch et al 2001[31]	high	moderate	moderate	High	high	N/A
Dewulf et al 2015[24]	moderate	moderate	low	moderate	moderate (treatment), high (other)	low
Green et al 2010[32]	high	moderate	moderate	low (infection), high (other)	moderate (infection), high (other)	low
Radtke et al 2017[33]	high	high	moderate	moderate	high	moderate
CF G-P Consortium	high	high	moderate	moderate	moderate	moderate

Study	Summary risk of bias by domain					
	Participation	Attrition	Genotype measure	Phenotype measure	Confounding	Statistical analysis
1993[34]						
Szczesniak et al 2017[36]	high	high	high	Low	moderate	low
de Boeck and Zolin 2017[37]	moderate	high	moderate	moderate	high	moderate
Dugueperoux de Braekeleer 2005[35]	moderate	low	low	moderate	moderate	high
Mackenzie et al 2017[23]	high	low	moderate	moderate	high	high

The majority of the included studies scored moderate to high risk of bias across the QUIPS domains. Much of the risk of bias related to lack of reporting of relevant information within the research papers.

The main strength of studies was that by using data from national CF registries or international consortiums the studies had information for several thousands of participants. This should give increased power for detecting differences in phenotype according to genotype. However, there were inherent limitations when using this collective data, which are discussed below.

Participation selection was at moderate or high risk of bias as many studies did not provide sufficient information about their participants or selection process. National registries would be expected to include the vast majority of people with CF from the countries or regions studied, yet only one study specified their patient coverage.[26] Additionally, studies did not explain the process by which patients are reported to the registries or how regularly their clinical data is entered. International consortiums typically represented less than half of the people with CF in the eligible countries or centres, for unclear reasons.[31, 33, 34]

Attrition bias was at moderate or high risk as there were high levels of missing data in follow-up assessments. Most studies had genotyping (and/or genotype classification data where relevant) available for between 50%[25, 28] and 75% of the full registry cohort.[24, 31] Studies applying further inclusion criteria, such as requirement for follow-up assessments, had data for even smaller subsamples.[27, 37] There is less potential for bias if

lack of genotyping, or initial patient entry into the registry or study, is random across all people with CF. However, there is concern for survivor bias in particular, where people with longer survival (and related genotypes) may be more likely to be genotyped and have phenotypic data available. This risk of bias relates not only to survival but to all outcome assessments. Alternatively, people with more severe disease (and related genotypes) may have more frequent clinic assessments and may be more likely to be genotyped and have their data entered into registries.

Moderate or high risk of bias for genotyping assessment related to a lack of information in the studies on how genotyping was performed and differences in classifications used in studies. As genotyping procedures were not described there is a risk that genotyping may have varied across centres and over time. In addition, as the functional effect and classification of variants is still ongoing and there is no definitive variant classification list (into classes 1-5), there were slight differences between studies in the groupings used for some classified variants. This may affect overall analyses comparing severe (class 1-3) with mild (class 4 or 5) variants.

Confounding is another key potential source of bias. Few studies adjusted for confounders and those that did varied in the factors they adjusted for. No studies reported the specific treatment or care received by patients (with the exception of one study specifically assessing treatment burden[24]), and no studies adjusted their analysis for any treatment received. However, some studies used geographic or temporal differences as crude proxies for treatment and care received. For example, one study (assessing survival) adjusted for birth year and size of treatment centre, another (assessing lung function decline) adjusted for cohort year, and 2 others (assessing multiple variables) compared age- and gender-matched F508del homozygotes and heterozygotes from the same centre.

The uncertain newborn screening context is another important confounder and no study adjusted for whether participants had been identified at birth through newborn screening. Some studies reported that newborn screening was not performed[23, 24, 28, 35] but for others this was unclear. Most studies pre-date the Millennium (and the births of many included participants would have been earlier) so would likely have been

conducted prior to the widespread implementation of newborn screening. However, there could be variability across US States and European countries in the timing of introduction.

Other studies carried out some adjustment for ethnicity[28, 32] or socioeconomic status.[28, 36] However, overall there was minimal adjustment for factors that may influence genotype-phenotype relationships.

Genotype association with each phenotypic outcome

Survival

Six studies assessed the link between *CFTR* genotype and survival.[25, 26, 28-30] Four were cohort studies comparing survival outcomes in individuals with different genotypes, and two were case control studies comparing the genotypes of older CF patients with the wider CF patient population. Overall, the studies showed a general association between class of the *CFTR* variant and survival outlook. However, it was not a precise correlation and there was a range of survival years for individuals carrying variants of the same functional class.

In 2003, in the largest registry study, McKone et al[26] found an association between survival and variant functional class when comparing F508del homozygotes with 11 common variants carried in heterozygosity with F508del. F508del homozygotes had a mortality rate about 20 times that of the general population (standardised for age and gender). The seven variants that conferred no difference in mortality compared with F508del homozygotes were also severe variants (class 1 to 3) like F508del (class 2). Three of the 4 variants with reduced mortality compared with F508del homozygotes were mild class 4 variants (see Table 3). The only disparate finding was reduced mortality for I507del heterozygotes, which is a class 2 variant like F508del.

In a 2006 follow-up study, McKone et al[25] compared variant functional classes and found that median survival of people carrying 2 severe variants (class 1 to 3) was considerably shorter than people carrying ≥1 mild variant (class 4 or 5), at 36 years compared with 50 years. This study was unique in showing that genotype is an independent predictor of survival adjusting for other phenotypic variables of lung function, infection, pancreatic sufficiency and nutritional status and cohort year and

treatment centre (a rough proxy for care/treatment). However, the authors found that variant functional class was not very accurate in predicting age at death. Using a cut-off of 30 years (the best combination of positive and negative predictive value, PPV and NPV), they found that there was very high sensitivity (98%) indicating that almost all people who die before age 30 years will have severe genotype (both variants class 1-3). However, the specificity was extremely low (11%) indicating that genotype would be an unreliable predictor of survival. The PPV of 69% suggests that around a third of people with severe genotype will live beyond the age of 30. Similarly the NPV of 71% shows that around a third of people with mild genotype (≥1 variant class 4 or 5) will die before age 30 years. Therefore, it is not possible to predict with certainty, the survival outlook for any individual with severe or mild variants.

A third study (Lai et al 2004)[27] also showed that F508del homozygotes are at risk of 'shorter' survival compared with people carrying mild class variants. However, this analysis was limited as the authors did not define what age range this meant.

The remaining studies compared F508del homozygotes with non-specific F508del heterozygotes.[28-30] Two of these supported poorer survival outlook for homozygotes, while one of two case-control studies did not find any difference in genotypes for people living above and below a set age cut-off (see Table 3). This inconsistency is likely, in part, due to the variability in genotypes among heterozygotes, which makes meaningful interpretation of these results difficult.

Overall almost all studies supported an association between genotype and survival. However, these associations were not strong enough for prediction. There were also limitations in the evidence. Firstly, survival outcomes were assessed in difference ways across studies (e.g. standardised mortality rates, predictive accuracy for mortality or risk of longer or shorter survival according to variant class, or comparing homozygotes and heterozygotes in people above and below age cut-offs), which precluded pooling of results. Secondly, studies did not describe how they identified patient deaths. There was no mention of accessing medical records or mortality registries. It is expected that deaths have been recorded in CF registries; however, it is difficult to judge whether records were complete and up-to-date. Finally, some studies differed in whether they counted transplant receipt as mortality[26,

[29] (based on the assumption that the patient would have died without transplant) while other studies did not state their approach to this issue.[25, 27, 28, 30]

Pancreatic status

Eight large registry studies compared pancreatic status in people with different variant class or specific genotype.[23, 24, 26, 31-35] Of all phenotypic outcomes, pancreatic status showed the most consistent association with variant class across studies. Compared with ≥1 class 4 or 5 variants, people with 2 class 1-3 variants, including F508del homozygotes, have a higher prevalence of pancreatic insufficiency. This was also almost universally found in the smaller single centre studies (Appendix 3, Table 28), despite their lower power for detecting differences.

Across the eight studies, between 90 and 100% of people with 2 severe class 1-3 variants, including F508del homozygotes, had pancreatic insufficiency and required enzyme replacement therapy. This was often associated with lower BMI, though differences in nutrition status were less consistently found. By contrast, people carrying ≥1 mild class 4 or 5 variants were comparatively less likely to have pancreatic insufficiency. However, many carrying mild class 4/5 variants could still have poor pancreatic function and nutrition status. Generally the larger registry studies[26, 31] found pancreatic insufficiency rates of 60-70% for those with at least one class 4/5 variant while the smaller registry studies and consortiums[24, 32-35] indicated lower insufficiency rates of around 25-50% in these groups.

The main limitation to pancreatic assessments is that all registry studies rely upon pre-collected and pre-recorded clinical data, usually collected across multiple centres. Most studies have considered patients to be pancreatic insufficient if use of enzyme replacement therapy (ERT) has been recorded in the registry. However, this may not be a precise indicator of the degree of insufficiency. The type of ERT, dose, frequency and duration of prescription may vary considerably between individuals across centres and between studies. Studies also did not report how frequently individual patient data was reported to registries or whether pancreatic status was a one-off status at patient entry. This is important as pancreatic insufficiency may develop or change over time.

Lung function

Ten studies[24, 26, 27, 31-37] assessed the association between genotype and lung function or risk of infection. Overall the association was weaker than that for pancreatic status and was less consistent within and across studies. Most studies found slightly better FEV1 (volume of air expired in the first second of forceful expiration), lower annual decline of FEV1 and lower rates of *Pseudomonas aeruginosa* colonisation in people carrying at least one mild class 4 or 5 variant compared with those carrying 2 severe class 1-3 variants.

Broadly studies demonstrated FEV1 of roughly 70-80% predicted for people with 2 severe variants and 80-90% for those with one or more mild variants. This 10% could make a clinically meaningful difference for people with milder genotype. However, there is some inconsistency and overlap in these ranges. For example, one European study[31] showed these ranges for patients under 18 years old but found that in patients over 18 years, FEV1 was lower than 70% for those with both mild and severe variants.

Other studies had inconsistent findings when analysing specific variants, finding that some mild class variants conferred improved lung function but not others. For example, McKone et al (2003)[26] found that compared with F508del homozygotes, F508del heterozygotes carrying mild class variants R117H, 2789+5G>A, and A455E had improved FEV1 and lower infection rates. However, the same was not found for mild class variants 3849+10kbC>T, R347P and R334W. Dugueperoux and De Braekeleer[35] similarly found that compared with F508del homozygotes, F508del heterozygotes carrying mild class 5 variant 2789+5G>A had improved FEV1, but not class 5 variant 3849+10kbC>T.

The smaller studies also had inconsistent findings. The majority found no difference in FEV1 when comparing variant classes, or F508del homozygotes vs heterozygotes, though they do have lower power to detect differences (see Appendix 3, Table 28).

The strength of this evidence is that lung function is expected to be recorded in a relatively standardised way by spirometry across centres while *P. aeruginosa* colonisation was most often assessed by looking at positive sputum cultures over a one-year period. However, as with

pancreatic status, it is not clear how consistently data may have been measured and entered into registries or reported to consortiums. Two studies[36, 37] carried out prospective assessments looking at decline in lung function over consecutive years or assessments and one analysed individuals with >1 follow-up assessment.[27] However, the remaining studies[24, 26, 31-35] did not clarify whether lung function has been averaged across multiple assessments for each individual or whether these were one-off measures.

Age at diagnosis

Five studies reported age at diagnosis for people of different variant class or genotype.[23, 24, 26, 34, 35] Age at diagnosis may serve as a general indicator of disease severity. All studies support a general pattern of infant/early childhood diagnosis for F508del homozygotes and heterozygotes carrying 2 severe class 1-3 variants while those with at least one class 4/5 variant are diagnosed at an 'older' age. However, this 'older' age could be highly variable from childhood through to adulthood.

Studies analysing by variant class found that people with mild class 4/5 variants are usually diagnosed at older age, which may indicate fewer symptoms and a milder disease course. However, the age at diagnosis by variant type is variable. One study[24] found that individuals carrying 2 severe variants were diagnosed by median 3 months of age compared with 5 years[24] for those carrying mild class 4 or 5 variants. However, the interquartile range for mild variants was very wide from a few months to 20 years, compared to a small range of only 1 month to 1 year for severe variants. Another study[26] found that people carrying 2 severe variants were diagnosed by mean 2 years compared with 11-12 years for people carrying mild variants.

Studies comparing F508del homozygotes with F508del heterozygotes[26, 34, 35] also followed this pattern of diagnosis aged around 2-3 years for people carrying F508del and another severe variant compared with late childhood or adolescence for people carrying a mild variant (see Table 6). However, McKone et al[26] also found older age at diagnosis for F508del heterozygotes carrying severe class variants G85E and I507del. This highlights the inconsistency within classes indicating that some individuals with 2 severe class variants may be diagnosed later in life. Of interest, one unclassified variant (P67L) was investigated. Genotype

F508del/P67L was associated with diagnosis in young adulthood at mean age 18 years.[23]

Most of the smaller studies comparing F508del homozygotes with F508del heterozygotes supported this association, finding that homozygotes were diagnosed in infancy and heterozygotes at older age (see Appendix 3, Table 28).[38-41] However, the mix of heterozygotes and variable design of these studies limits interpretation (for example, one compared genotypes of those diagnosed before or after 6 months, another compared those diagnosed early or in late adulthood).

Age at diagnosis may be expected to be consistently reported across studies and centres. However, this could encompass variable methods of presentation for individuals, such as by clinical symptoms, family history or screening. The uncertain newborn screening context is a notable limitation. Four of the studies reported that newborn screening was not performed[23, 24, 28, 35] but this is unclear for the large US registry study[26] and European consortium.[34] As assessment periods for these 2 studies were during the 1990s, most individuals are expected to have been born prior to the widespread implementation of newborn screening. However, there may have been variability within US states and across European countries in the timing of introduction.

Treatment burden

Dewulf et al[24] was the only study to have compared treatment burden between variant classes. This assessment supports the general theme of all other findings by phenotype. People carrying class 4 or 5 variants needed fewer and less intense treatments such as intravenous antibiotics or parenteral nutrition than people carrying two severe class 1-3 variants. They were also less likely to be hospitalised over the course of one year. However, little can be concluded from this single study and other studies would be needed to confirm this association.

Overall interpretation

The various quality limitations around representation, lack of genotyping, variable phenotype assessment and lack of adjustment for confounders were fairly consistent across studies. Despite these weaknesses, there was general replication of findings across studies, indicating that there is a definite relationship between genotype and phenotype. At least one

class 4 or 5 variant appears to confer milder disease course even in the presence of a severe class 1-3 variant (typically F508del), whereas nearly all people carrying 2 severe class (1-3) variants have more severe disease. However, there is wide variation among individuals with variants in the same functional class (1-3 or 4-5) and not all individuals will follow the same pattern. Therefore, there would be a need for caution if using fetal variants alone to support informed decision-making and guide pregnancy decisions, as the estimation of phenotypic outcomes is not precise. It would be difficult to predict with any certainty how an individual's clinical disease is likely to progress or what their life expectancy could be.

It would be possible to say that nearly all people carrying 2 severe class (1-3) variants, particularly F508del homozygotes, will have pancreatic insufficiency (the clearest and most consistent genotype-phenotype association) and will be at higher risk of early mortality. They may also have lower lung function and earlier age at diagnosis. Similarly people carrying at least one class 4/5 variant are more likely to have pancreatic sufficiency, a relatively good survival outlook and may also have better lung function and later diagnosis. However, there is wide variation among individuals with variants in the same functional class. Variant class could give a rough guide of survival outlook, but it would not be possible to predict life expectancy with any accuracy. Though most people who die before 30 years have 2 severe class 1-3 variants, around a third with these genotypes may live beyond this age. Similarly, though most people carrying at least one mild class 4/5 variant would be expected to live beyond 30, around a third could die before this age. Likewise, anywhere between one- and two-thirds of people with mild class could be pancreatic insufficient. Lung function (FEV1) may be around 70-80% predicted for people with 2 severe class variants and around 80-90% for those with at least one mild, but these are only broad estimates and were inconsistent across studies. Disease outlook in terms of need for treatment, clinic visits or hospitalisation and overall quality of life would be very difficult to predict based on the available evidence.

In addition to this uncertainty, survival rates have also improved for people with CF in past decades. This is likely due to improved care and treatment. In particular, ivacaftor now offers improved outlook for people with class 3 variants, and there could be new treatment advances in the future. Therefore, the overall clinical prognosis over the coming decades

could vary for individuals born today with any genotype, limiting the applicability of these findings.

It is also difficult to know from this evidence which variants should be included in potential antenatal screening panels. The current ACMG panel[5] of 23 variants (question 3, Table 12) includes the class 4 variant R117H, which is the 3rd most common variant among people with CF in the UK in 2016.[1] The gathered evidence consistently indicates milder disease course with this variant. The panel also includes mild class 5 variants 3849+10kbC→T, 2789+5G→A and A455E, and rarer class 4 variants R347P and R334W.

As further discussed in question 3, a 2003 UK study[4] proposed modifications to the ACMG panel suggesting removal of R117H and 3849+10kbC→T on the basis of milder disease course, and the addition of the class 1 variant 1078delT (ACMG had included this in their original panel but removed it due to population frequency <0.1%). However, this proposed panel still retains 2789+5G→A, A455E, R347P and R334W. The study authors do not report the rationale for retaining these class 4/5 variants, for example whether based on additional evidence or clinical experience. There could be a case for excluding all class 4/5 variants from potential antenatal screening panels. However, this review suggests that the evidence on genotype-phenotype correlation may not yet be strong enough to make this decision.

If variant functional class is used as a basis for predicting phenotype, this raises questions related to unclassified variants that have not been included in previous antenatal screening panels. This includes, for example, P67L, D1152H, Q493X or 3272-26A→G each carried by around 1% of people with CF in the UK.[1] MacKenzie et al[23] studied P67L, specifically, and found it was associated with late diagnosis in early adulthood and pancreatic sufficiency. McKone et al[26] also demonstrated that as a pooled group all unclassified variants were associated with milder disease course with pancreatic sufficiency, improved lung function and later diagnosis compared with F508del homozygotes (Tables 4-6).

However, it is unknown what effect all unclassified, and previously unscreened variants, may have. While many could cause only mild disease, single organ involvement or CFTR-related disorders, there could be variability. For example, D1152H is said to have been seen in both

classic CF and non-classic CF and related disorders.[9] Some recent studies have now added this to the group of class 4 variants.[24, 32, 33] This shows that understanding around the functional and phenotypic effects of variants is still developing.

With so many as yet unclassified *CFTR* variants it is difficult to know which could cause only mild disease and which may have severe effect. Antenatal CF screening could potentially give couples 'false reassurance' that they will never have a child with CF. Screening for the common and classified variants could give extremely high reassurance of this (as discussed in question 3, criteria 4 and 8 below). However, there is the very small risk that rarer unclassified variants that are not screened for could cause severe CF.

Summary of Findings Relevant to Criterion 1 – genotype-phenotype association: Criterion not met[§]

There is relatively consistent evidence from 15 studies that functional class of the disease-causing gene variant(s) is associated with phenotype in CF. However, phenotype is variable for individuals with the same genotypes, which means that clinical disease course cannot be accurately predicted from an individual's genotype.

F508del homozygotes and people carrying 2 severe class 1-3 variants are likely to have pancreatic insufficiency and poorer survival outlook. Comparatively, carriage of at least one mild class 4 or 5 variant (such as the common variant R117H) confers milder disease course as evident by lower rates of pancreatic insufficiency and longer survival. There are similar, though less consistent, associations with lung function and age at diagnosis. Only one study looked at the relationship between genotype and treatment burden, and no studies looked at the relationship with quality of life. These outcomes are likely to be important for patients and their families.

Despite this, phenotype is highly variable for individuals with variants in the same functional class or with the same genotype. Across the studies, up to 10% of people with 2 severe class variants (including F508del homozygotes) had pancreatic sufficiency while 25-70% of people carrying at least one mild class variant had pancreatic insufficiency. Likewise, most people who died before the age of 30 years had severe variants, but a third with mild variants also died before this age, while a third with severe variants survived longer. Carriage of at least one mild class variant was usually associated with improved lung function, but this was not consistently seen across all studies. Similarly diagnosis could be made at any time from childhood to adulthood for people carrying at least

§ §**Met** -for example, this should be applied in circumstances in which there is a sufficient volume of evidence of sufficient quality to judge an outcome or effect which is unlikely to be changed by further research or systematic review.
Not Met - for example, this should be applied in circumstances where there is insufficient evidence to clearly judge an outcome or effect or where there is sufficient evidence of poor performance.
Uncertain -for example, this should be applied in circumstances in which the constraints of an evidence summary prevent a reliable answer to the question. An example of this may be when the need for a systematic review and meta-analysis is identified by the rapid review.

one mild class variant. This variability would make it difficult to predict disease course with any certainty, which in turn would complicate counselling on pregnancy decisions.

Another issue is the large number of potentially causative variants for CF. Only a small number of the more common variants have been widely studied, classified and included in antenatal screening panels to date. Though many of the rarer, non-screened variants may have mild functional effect, this cannot be said with certainty.

The 15 included studies also have common limitations in their quality and applicability. Most were registry-based and lacked genotyping or variant classification for around half of the registry population. Therefore results may not be representative of the CF population as a whole. Few studies adjusted for confounders, in particular treatment and newborn screening, increasing the potential for bias. Most studies also date from over 20 years ago and may not be applicable to the treatment context today. New disease-specific treatments have altered the prognosis for people with certain genotypes, and survival is continually improving, which is likely to be a reflection of care and treatment improvements in general.

Overall, there is a consistent association between genotype and phenotype. However, the variability in outcomes for individuals, moderate to high risk of bias across studies (particularly relating to attrition and confounding), limited applicability to care today, and uncertain effects of rare variants mean that this part of the criterion is not met. The available evidence indicates that it is not possible to use genotype to predict phenotype with sufficient accuracy to allow pregnant women/couples identified through antenatal screening to make fully-informed reproductive decisions. This degree of uncertainty is considered a reasonable price to pay in newborn screening as more babies will benefit than be harmed from screening. However, in antenatal testing where the option is to continue or terminate the pregnancy, a much higher degree of certainty is needed. Furthermore, in the newborn programme, mutation analysis is a second step only carried out for infants with immunoreactive trypsinogen levels above the cut-off (on 2 assays).

Criterion 4 – Test accuracy

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

Criterion 8 – Mutation selection

If the test is for a particular mutation or set of genetic variants the method for their selection and the means through which these will be kept under review in the programme should be clearly set out.

Question 3 – To describe the genotypes/mutations covered by the commercially available tests for antenatal CF screening in the UK. To estimate whether these tests are clinically accurate for diagnosing CF in the fetus or newborn.

Background

The Murray et al HTA[3] concluded that antenatal screening for CF was feasible and could be offered routinely to women and their partners in all maternity units. This followed the publication of 11 studies of antenatal screening pilots, 5 of which were conducted in the UK. A summary of these UK studies as reported by Murray et al[3] is presented in the Table 9 below. The remaining six pilots were conducted in the US, Germany and Denmark.

Studies predominantly performed stepwise screening, where the mother is tested and only if she is a carrier is the partner invited for screening. With stepwise screening the individual becomes aware of their carrier status. Some studies instead performed couples screening where both parents were tested, either with disclosure (couples told their individual results) or non-disclosure (informed of positive or negative carrier status as a couple) of the results.

Murray et al[3] reported that across all 11 pilots, 50,801 women were invited for screening with a pooled uptake rate of 74%. When following stepwise screening, 92% of fathers received testing if the mother was identified to be a carrier. Invasive diagnostic antenatal testing was performed for 89% of all carrier couples. In all but one case where the baby was found to carry 2 CF disease-causing variants (17/18, 94%), the pregnancy was terminated.

The studies revealed similar uptake between stepwise and couples testing. The test in the Scottish, Leeds and Manchester studies was able to detect 86% of known variants and gave a carrier frequency of 1 in 28. This would equate to an overall carrier frequency in the UK of around 1 in 24 (if a test could detect all variants). The overall false positive rate (among carriers) in these studies was reported at 0.1%.

The studies did not report problems, and antenatal screening seemed feasible from a practical perspective, setting aside other psychological and ethical aspects (as addressed by question 4).

This review question therefore aimed to look at whether there is further evidence on the clinical test accuracy of antenatal screening tests for CF. That is, the accuracy of tests to predict CF diagnosis in the fetus or newborn. The aim was to see which variants had been covered in such commercially available tests in the UK. The purpose was not to look at the analytical validity of these tests to detect the intended panel of variants.

Eligibility for inclusion in the review

Eligible studies would be any cohort or pilot studies of antenatal screening programmes that had been published since 2000. Studies could be either from the UK or alternatively from similar Western populations where the tested variant panel may be applicable.

Of particular interest were any studies that had comprehensive follow-up for both screen positive and negative couples; for example, seeing whether any child born to screen-negative couples developed CF. This is something lacking from prior screening pilots, and would allow calculation of sensitivity, specificity and negative predictive value (NPV) of a given test to detect CF in the fetus/newborn.

The review aimed to consider either couples or stepwise screening in non-selected samples of pregnant women (for example, random or consecutively enrolled) who would represent the general pregnant population. Data by ethnicity would be reviewed if this was available.

The review did not intend to cover screening of high risk couples, such as those with family history or previous pregnancy with CF.

Table 9. Summary of UK antenatal screening pilots 1990s as reported by Murray et al[3]

Location	Variants tested	Reported variant coverage for that region	Sample collected	Strategy	Screening uptake among women	Partners of carriers tested	Antenatal diagnosis performed for carrier couples	Termination of CF pregnancies identified
Edinburgh 1992-94	F508del, G542X, G551D, 621+1G→T* plus R553X, 1105del	85%	Blood and mouthwash	Stepwise Couple	4978/6030 (83%) 12,566/16,571 (76%)	189/190 (99%) NA	33/36	13/13
Aberdeen 1995	F508del, G542X, G551D, 621+1G→T*	92%	Mouthwash	Stepwise Couple	1487/1641 (91%) 321/361 (89%)	47/48 (98%) NA	2/2	0/0
Leeds 1993	F508del	80-90%	Blood	Stepwise	3773/6071 (62%)	127/130 (98%)	1/3	0/0
Manchester 1995	F508del, G542X, G551D, 621+1G→T* plus W1282X	CF4: 85% W1282X to cover Ashkenazi Jewish	Mouthwash	Mixed (stepwise and couple: psychological aspects assessed)	529/623 (85%)	10/10 (100%)	1/1	0/0
Oxford 1993	F508del, G542X, G551D, 621+1G→T* plus R553X, W1383X, R1283M	Unreported (initially F508del, G551D R553X then extended)	Buccal smear	Couple	543/810 (67%)	NA	0/0	0/0

*CF4 group of variants tested using Cellmark Diagnostics kit. Additional variants tested using local in-house assays.

Description of the evidence

Forty-seven studies from the search were considered relevant to this question at initial appraisal, of which 12 were selected for full text appraisal.

Only a single Australian study[42] of an antenatal CF screening pilot has been published. This study is summarised in Table 10, with full evidence extraction presented in the summary and appraisal of individual studies in **Error! Reference source not found.**, Table 29.

No other studies met inclusion criteria to provide evidence for this question. Two narrative reviews present the position of opinion around the panel of variants to include in antenatal screening in the UK[4] and US[5] and the likelihood of false negatives with these tests. These studies do not provide evidence for this question, but are discussed below as they give useful background to the situation.

The remaining 9 studies reviewed at full text were excluded. Most excluded studies concerned either screening of high-risk couples or preimplantation genetic diagnosis for couples receiving assisted conception. Appendix 2 lists the studies excluded at full text appraisal for this question, with the reason for exclusion.

Discussion of findings

Findings and critical appraisal of the Australian screening pilot

The single Australian study demonstrates the practical experience of antenatal CF screening performed over the past 18 years.

It screened 3200 individuals and gave a carrier frequency of 1 in 30. If the test covers 84% of variants, this roughly equates to a carrier rate of around 1 in 25 in the general population. Six carrier couples were identified, all of whom accepted diagnostic testing and 2 fetuses were found to carry 2 CF variants (see Table 10). This gives a PPV of 33% for a positive couples-screening test to indicate CF in the fetus. This is in general agreement with a carrier couple having a 1 in 4 probability of having an affected child. The PPV will be influenced by the prevalence of variants in the population.

Both affected pregnancies were terminated, as has been the predominant experience with past screening pilots. However, it is not possible to know how severely affected by CF these infants may or may not have been.

There was no further pregnancy or birth follow-up for the cohort. Therefore it is not possible to calculate sensitivity, specificity or NPV of the test or know how many false negatives for CF may have resulted from other variants not covered by the panel. False positives for parental carriage of these variants are unlikely but again cannot be assessed from this study.

The study has limited applicability to the UK. Firstly as participants were required to pay for the test it may not represent the general pregnant population. For example, participating women/couples may have higher socioeconomic status than non-participants. Their carrier frequency may differ. Secondly, the panel of variants was selected to give good coverage of the local Australian population. However, it differs from the panel used in previous UK pilots and the most prevalent variants in the UK (see Table 11). The results from this study could not inform what would be seen if this same variant panel was used in the UK.

Table 10: Screening pilot, Australia

Location	Screening strategy	Variant tested	Uptake	Carriers	Carrier couples	Outcome
Massie et al 2009[42] Victoria, Australia 2006-08	Pay-for test offered to women or couples attending a GP: • Prior to pregnancy • <14 weeks pregnant Couples screening recommended but mostly stepwise. Method: check swab.	12 variant panel covering 83.5% of the general population of the region and 95% of the Ashkenazi Jewish population	Total 3200 screened: • 3000 women • 200 men Including 100 couples (200 individuals)	106 carriers detected: • 92 women • 14 men Frequency 1 in 30 None part of couples screening: • 106 partners tested	9 carrier couples: 3 pre-conception 6 pregnant	6/6 pregnant couples accepted CVS: 2/6 affected fetuses (PPV 33%) Both terminated. No follow-up of screen negatives.

Variant panels that may be used for antenatal screening in the UK – summary of narrative reviews

It is not known what variant panel would be used for universal antenatal screening in the UK. The UK Genetic Testing Network[43] currently lists 19 laboratories that offer antenatal testing. Targeted mutation analysis is the most common service available where the test would be for a select panel of variants (those tested not given). Other laboratories provide testing for known variants carried by family members or gene tracking.

The narrative review by Wald et al[4] (2003) was the only post-2000 publication that considered antenatal screening from the UK perspective. Wald et al summarise the theoretical probabilities of having an affected pregnancy with couples-screening based on a carrier frequency 1 in 25 (birth prevalence 1 in 2500) and a test that identifies 85% of carriers (Table 11).

Table 11: Theoretical screening, Wald et al[4]

	250,000 couples screened		
Scenario	A	B	C
Parent carrier status	+/+	+/-	-/-
Pregnancies affected	288	8211	233,290
Fetus without CF	216	8198	233,288
Fetus with CF	72	13	2
Probability of fetal CF	1 in 4	1 in 632	1 in 116,645

Wald et al estimate that a test that could identify 85% of variants would miss only 1 in 9000 CF-affected pregnancies among white Europeans. The proportion in other ethnicities (based on pre-2000 prevalence estimates) would be expected to be even lower: 1 in 14,000 among Asians and 1 in 20,000 among Afro-Caribbean.

In 2001 the American College of Medical Genetics (ACMG) established a 25 variant panel including variants with a carrier frequency of >0.1% in the US population. Wald et al proposed a revised 22 variant panel that they suggested could identify 85% of variants in the UK. This involved removal of 3 variants from the ACMG panel that they considered to have

low penetrance (I148T) or to be associated with less severe disease (R117H and 3849+10kbC→T).

Brennan et al[5] (2016) report that the ACMG 25 variant panel was revised in 2004 removing 2 variants with low penetrance (I148T and the class 1 variant 1078delT, which was retained by the suggested panel of Wald et al). The ACMG panel did not, however, remove mild class variants. Table 12 presents the variants covered by each of these proposed screening panels, and those used in the screening pilots, against variant frequency in the 2016 UK CF register.

Brennan et al estimated that the probability of having an affected pregnancy with negative couples-screen on this revised panel was less than 1 in 40,000 for white Americans, lower for other ethnicities.

Overall the theoretical data suggests that the likelihood of false negatives from antenatal screening would be low. However, the inconsistency in variant panels, both in theory and in practice, indicates that as yet it is uncertain which variants would be included in an antenatal screening test in the UK. Even if the test aimed to cover the most frequent variants, there could be no certainty that screen-negative couples would not have a child with CF due to less common variants. Test performance could also vary regionally within the UK depending on the population demographic.

There is also the question of whether couple or stepwise screening would be carried out in the UK. Wald et al consider that couples screening would be preferable. This would designate a positive result only if both couples screen negative so avoiding the scenario of one person being denoted a variant carrier. However, both the Massie et al study and past screening pilots highlight some feasibility issues with couples-screening as the uptake is low.

Table 12: Variants prevalent in the 2016 UK population with CF, and those covered by screen tests and pilots

							Pre-2000 UK screening pilots[3]				
Variant (Legacy name)	**Human Genome Variation Society (HGVS) name (nucleotide)**	**Class of variant in panel#**	**UK 2016 % with CF carrying ≥1 variant[1]**	**ACMG 2004 panel[5]**	**Wald et al[4] theoretical panel for UK**	**Massie et al[42] 2006-08**	**Edinburgh 1992-94**	**Aberdeen 1995**	**Leeds 1993**	**Manchester 1995**	**Oxford 1993**
F508del	c.1521_1523delCTT	2	90.9	♦	♦	♦	♦	♦	♦	♦	♦
G551D	c.1652G→A	3	5.9	♦	♦	♦	♦	♦		♦	♦
R117H	c.350G→A	4 (mild)	5.1	♦	removed*						
G542X	c.1624G→T	1	3.6	♦	♦	♦	♦	♦		♦	♦
621+1G→T	c.489+1G→T	1	2.6	♦	♦		♦	♦		♦	♦
N1303K	c.3909C→G	2	1.7	♦	♦	♦					
1717-1G→A	c.1585-1G→A	1	1.4	♦	♦						
1898+1G→A	c.1766+1G→A	unknown	1.3	♦	♦						
3659delC	c.3528delC	1	1.1	♦	♦						
P67L	c.200C→T		1.1								
D1152H	c.3454G→C		1.0								
R560T	c.1679G→C	2	1.0	♦	♦	♦					
I507del	c.1519_1521delATC	2	0.9	♦	♦	♦					
Q493X	c.1477C→T		0.9								
3272-26A→G	c.3140-26A→G		0.8								

Variant (Legacy name)	Human Genome Variation Society (HGVS) name (nucleotide)	Class of variant in panel#	UK 2016 % with CF carrying ≥1 variant[1]	ACMG 2004 panel[5]	Wald et al[4] theoretical panel for UK	Massie et al[42] 2006-08	Pre-2000 UK screening pilots[3]				
							Edinburgh 1992-94	Aberdeen 1995	Leeds 1993	Manchester 1995	Oxford 1993
R553X	c.1657C→T	1	0.8	♦	♦	♦	♦				♦
G85E	c.254G→A	2	0.8	♦	♦						
3849+10kb C→T	c.3717+12191 C→T	5 (mild)	0.8	♦	removed*						
E60X	c.178G→T		0.7								
1154insTC	c.1022_1023insTC		0.6								
W1282X	c.3846G→A	1	0.6	♦	♦	♦				♦	
2789+5G→A	c.2657+5G→A	5 (mild)	0.5	♦	♦						
1078delT	c.948delT	1	0.5		♦ retained*						
S549N	c.1646G→A	3	0.4								
2184delA	c.2052delA	unknown	0.4	♦	♦						
R347P	c.1040G→C	4 (mild)	0.4	♦	♦						
A455E	c.1364C→A	5 (mild)	0.4	♦	♦						
L206W	c.617T→G		0.3								
R1162X	c.3484C→T	1	0.3	♦	♦						
V520F	c.1558G→T	not stated	0.3			♦					
711+3A→G	c.579+3A→G		0.3								

Variant (Legacy name)	Human Genome Variation Society (HGVS) name (nucleotide)	Class of variant in panel#	UK 2016 % with CF carrying ≥1 variant[1]	ACMG 2004 panel[5]	Wald et al[4] theoretical panel for UK	Massie et al[42] 2006-08	Pre-2000 UK screening pilots[3]				
							Edinburgh 1992-94	Aberdeen 1995	Leeds 1993	Manchester 1995	Oxford 1993
5T	c.1210-12[5](AJ574948.1:g.152T[5])		0.3								
2789+2insA	c.2657+2_2657+3insA		0.2								
3120+1G→A	c.2988+1G→A	not stated	0.2	♦	♦						
R352Q	c.1055G→A		0.2								
R347H	c.1040G→A		0.2								
E585X	c.1753→T		0.2								
2711delT	c.2583delT		0.2								
R334W	c.1000C→T	4 (mild)	0.2	♦	♦						
1525-1G→A	1393-1G→A		0.2								
R1158X	c.3472C→T		0.1								
S945L	c.2834C→T		0.1								
G178R	c.532G→A	3	0.1								
Y569D	c.1705T→G		0.1								
R709X	c.2125C→T		0.1								
2184insA	c.2052_2053insA		0.1								

							Pre-2000 UK screening pilots[3]				
Variant (Legacy name)	**Human Genome Variation Society (HGVS) name (nucleotide)**	**Class of variant in panel#**	**UK 2016 % with CF carrying ≥1 variant[1]**	**ACMG 2004 panel[5]**	**Wald et al[4] theoretical panel for UK**	**Massie et al[42] 2006-08**	**Edinburgh 1992-94**	**Aberdeen 1995**	**Leeds 1993**	**Manchester 1995**	**Oxford 1993**
R1066C	c.3196C→T		0.1								
711+1G→T	c.579+1G→T	unknown	0.1	♦	♦						
R1066H	c.3197G→A		0.1								
S489X	c.1466C→A		0.1								
S1235R	c.3705T→G		0.1								
1811+1G→C	c.1679+1G→C		0.1								
R117C	c.349C→T		0.1								
Q98X	c.292C→T		0.1								
A559T	c.1675G→A		0.1								
R75X	c.223C→T		0.1								
R75Q	c.224G→A		0.1								
S549R	c.1645A→C	3	0.1								
K710X	c.2128A→T		0.1								
Other		not stated	14.1			♦ 1585-1G→A 489+1G→T 3718-	♦ 1105del				♦ W1383X R1283M

Variant (Legacy name)	Human Genome Variation Society (HGVS) name (nucleotide)	Class of variant in panel#	UK 2016 % with CF carrying ≥1 variant[1]	ACMG 2004 panel[5]	Wald et al[4] theoretical panel for UK	Massie et al[42] 2006-08	Pre-2000 UK screening pilots[3]				
							Edinburgh 1992-94	Aberdeen 1995	Leeds 1993	Manchester 1995	Oxford 1993
						2477C →T					

*removal or retention proposed by Wald et al differentiating from ACMG panel

#classification based on de Boeck et al (2014)[44] and McKone et al (2006)[26]

Summary of Findings Relevant to Criteria 4 and 8: Criterion not met**

The literature indicates there is no well-established mutation or variant panel that could be used in an antenatal screening test for CF in the UK. Most suggested tests aim to identify ≥85% of carrier variants in the given population to minimise the chance of screen-negative couples having a pregnancy affected by CF. However, there is inconsistency and uncertainty over the panel of variants to include. Pre-2000 UK pilots have included the CF4 panel of variants (F508del, G542X, G551D, 621+1G→T) which are the most prevalent variants in the UK (and also tested in newborn screening), but there is high variability among others included.

Only a single Australian screening pilot has been published since 2000. The PPV (fetus with 2 variants if both parents were carriers) was calculated at 33%. Consistent with pre-2000 screening pilots, antenatal diagnosis resulted in termination of both affected pregnancies. Birth outcomes were not followed for screen negative couples so it is not known if any screen-negative parents may have had a child affected by CF due to other variants.

The findings of this Australian study have limited applicability to the UK. The test panel was based on variants frequent in the local population of Victoria. It excluded several variants common in the UK and included other rarer ones. Furthermore the programme included preconception in addition to antenatal screening. It was also a pay-for screening service which limits representation and may increase bias.

As a result of the low volume of evidence, limited applicability and risk of bias, this criterion is not met.

** **Met** -for example, this should be applied in circumstances in which there is a sufficient volume of evidence of sufficient quality to judge an outcome or effect which is unlikely to be changed by further research or systematic review.
Not Met - for example, this should be applied in circumstances where there is insufficient evidence to clearly judge an outcome or effect or where there is sufficient evidence of poor performance.
Uncertain -for example, this should be applied in circumstances in which the constraints of an evidence summary prevent a reliable answer to the question. An example of this may be when the need for a systematic review and meta-analysis is identified by the rapid review.

Criterion 12 – Acceptability of screening

There should be evidence that the complete screening programme (test, diagnostic procedures, treatment/ intervention) is clinically, socially and ethically acceptable to health professionals and the public.

Question 4 – Is an antenatal screening programme acceptable to people in the UK: pregnant couples, individuals with CF carrier status or individuals with CF?

Background

The Murray et al HTA[3] revealed several potential psychological and emotional issues around antenatal screening for CF, including:

- many couples with negative results may falsely believe they have no risk of having a child affected by CF
- some people experience anxiety from the screening process
- people identified as carriers may experience stigmatisation
- prenatal diagnosis (amniocentesis or CVS) carries risk of miscarriage that it is difficult to quantify
- views on antenatal screening among people affected by CF are rarely obtained: past surveys found preference or acceptance for preconception and newborn screening but only half found termination of an affected pregnancy acceptable

Acceptability of an antenatal CF screening programme among the target population is of paramount importance. The results would be used to guide pregnancy decisions, and most screening pilots have reported that nearly all couples with a positive antenatal diagnosis terminate the pregnancy. A screening programme would be expected to reduce the prevalence of CF.

It is important to understand the views and feelings of the general population who would be screened, people with positive results, and the views of people living with CF. This review aimed to assess the acceptability of antenatal CF screening in the UK population.

Eligibility for inclusion in the review

Populations of interest were the general UK population, pregnant women or couples who could be invited for screening, people with CF carrier status, and people with CF or those affected by CF (for example, family members). The aim was to look at views on a universal antenatal screening programme which allowed detection of a fetus carrying 2 CF variants for the purpose of informed decision-making (whether to continue with or terminate the pregnancy).

Eligible study designs were broad, including pilot or feasibility studies, cohort studies, cross sectional studies or qualitative studies (such as focus groups or interviews) including more than 10 participants. Only studies from the UK were eligible for inclusion in order to ensure cultural and sociodemographic representation of this population.

Description of the evidence

Twenty-nine studies from the search were considered potentially relevant to this question at first pass appraisal, 10 of which were selected for full text appraisal. No studies met the inclusion criteria.

No UK studies were identified that assessed views on universal antenatal screening programmes among the general population, people with carrier status or people affected by CF.

Studies assessing only views on preconception screening only, rather than antenatal screening, were excluded. Several studies (including one UK study[45]) included people who previously had a child with CF but were excluded as they only questioned their own subsequent reproductive decisions rather than their views on antenatal screening in general.

Another UK study[46] questioned people affected by CF about the aspects of the condition they consider most important to provide information on in antenatal screening and diagnostic testing materials. For example, “people with CF can die at young age” or “males are almost always infertile”. This was in the context of cascade or high-risk testing as currently provided in the UK. The study did not solicit any views on whether screening, diagnostic testing or termination seemed acceptable

to them or not, either in this context or when considering universal antenatal screening.

In summary, no studies were identified that assessed views on universal antenatal screening programmes among the general population, people with carrier status or people affected by CF in the UK. Therefore, there is no evidence for this criterion.

Non-UK studies identified by the search

The search identified 3 non-UK studies. Two assessed the views of people participating in the Australian antenatal screening pilot reported by Massie et al.[42] One [47] considered the views of individuals with positive carrier status (though with a negative partner) and those of a random sample of screen-negatives. The second[48] considered the effect on screen-positive couples. A third Belgian study[49] questioned views of people with CF or their parents about general population screening.

As there were no studies from the UK, a brief summary of the general themes emerging from these 3 studies is given. However, this is not intended to be a comprehensive review of the international evidence nor does it constitute evidence for this criterion. The search was targeted to identify UK studies and there may be additional non-UK evidence. Furthermore views are culturally representative and the views and opinions from these studies are not applicable or generalisable to the UK population. Instead, the summary is given to provide some context of some of the views from different countries.

Discussion of findings

Views of participants of the Australian antenatal screening pilot

One of the key themes that emerged from the 2 Australian studies is the limited prior understanding about the purpose and implications of CF screening.

Ioannou et al (2010)[47] mailed questionnaires to a random sample of 162 non-carriers, and 79 people found to be carriers but who had screen-negative partners (therefore no pregnancy decision was needed).

Questionnaire response rates were 62% for carriers and 46% for non-carriers.

The main reason for participating in screening was the belief that CF was a severe condition. Most understood their carrier test result, but there was misunderstanding about what this meant for them. Over half incorrectly thought that a CF test can identify all carriers and that if no CF variant was identified this meant they could not be a carrier. Carriers reported no greater anxiety than non-carriers, but this may have been different had they also had a carrier partner.

A later study by Ioannou et al (2015)[48] questioned the psychological impact and decision making of carrier couples (where both partners screened positive) identified during the complete pilot (2006-12). Six of 9 carrier couples were pregnant when screened (the study also included preconception screening).

Most had minimal knowledge of CF before screening but understood the implication of their result through counselling. All 6 couples opted for antenatal diagnosis, something that was not even viewed as “a decision” but simply the next step in the process. Most reported high levels of anxiety while waiting for the result. The 2/6 couples with a positive antenatal diagnosis reported devastation and grief. Both terminated: one couple had already decided that was what they would do if they got a positive result, but the other had difficulty and uncertainty coming to that decision.

Looking at future effects, 3/9 carrier couples had no further pregnancies and the 4/6 who did, chose antenatal diagnosis. One of these pregnancies was affected and the parents found the decision to terminate very traumatic and regretted the decision afterwards. With a subsequent pregnancy they decided to keep the baby even if they had CF. Similarly one of the couples who had terminated the baby during initial screening faced a dilemma over whether to have antenatal diagnosis a second time.

This study also reported the additional complexity of couples reporting their carrier status to other family members who did not understand what this could mean for them.

As discussed in question 3, the Australian screening pilot had limited applicability to potential UK screening programmes as it was a pay-for service and included preconception screening. With this is mind, nearly all carriers identified in the former Ioannou et al (2010)[47] study, believed that screening should be offered before pregnancy.

Views of people with CF or their parents

One Belgian study (Janssens et al 2016)[49] sent questionnaires to 134 people with CF (or their parents) about their views on carrier screening. Sixty percent of the sample was Catholic so their views may not be representative of the general population.

The majority of those questioned (80%) believed population-based screening could offer more benefits than harms, but there were reservations. Half believed that it would increase the number of pregnancy terminations and nearly a third felt that it would cause less investment in treatments for CF.

Nevertheless almost all participants (96%) did not feel that screening should be limited only to those with a family history of CF. Antenatal screening was acceptable to 73%, though preconception screening was the preferred option, accepted by 86%.

Summary of Findings Relevant to Criterion 12: not met††

No studies have assessed views on universal antenatal CF screening among the UK population. Therefore this criterion is not met.

The search identified several studies from non-UK populations. Studies from the Australian screening pilot suggest some lack of understanding around carrier status, for example thinking a negative test means you cannot be a carrier. All couples screening positive opted for antenatal diagnosis (in current and subsequent pregnancies) but experienced high anxiety waiting for the result. All 3 positive pregnancies were terminated but 2/3 sets of couples reported high levels of grief and regret around the decision.

One Belgium study assessed people affected by CF (majority Catholic). Most thought that population-based screening seemed acceptable, but there were concerns it would detract resources from CF and increase termination rates. Preconception screening was the preferred option.

These studies do not represent all of the international literature on screening views and are culturally-specific so cannot be generalised to the UK.

†† **Met** -for example, this should be applied in circumstances in which there is a sufficient volume of evidence of sufficient quality to judge an outcome or effect which is unlikely to be changed by further research or systematic review.
Not Met - for example, this should be applied in circumstances where there is insufficient evidence to clearly judge an outcome or effect or where there is sufficient evidence of poor performance.
Uncertain -for example, this should be applied in circumstances in which the constraints of an evidence summary prevent a reliable answer to the question. An example of this may be when the need for a systematic review and meta-analysis is identified by the rapid review.

Addendum to Criterion 12

Appraisal of this external review by the Fetal Maternal Child Health (FMCH) group highlighted one additional paper relevant to the acceptability of antenatal CF screening in the UK, published after the search date of this review. The Boardman and Hale (2018) publication assessed the views on 'selective reproduction' of people affected by different genetic conditions. Thematic analysis was used on qualitative interview data, and participants were categorised as supporting, not-supporting, or having ambivalent views toward selective reproduction. The findings related only to the participants with CF are summarised in Table 13 and below.

Table 13: Post-search publication on acceptability

Study	Design	Population	Interview questions	Views
Boardman and Hale 2018[50] **UK**	Qualitative interviews conducted March 2017 to 2018. Part of a study assessing views of people with CF and 3 other conditions (haemophilia, thalassaemia and fragile X).	N=10* adults with CF (50% female, age range 21-58, 40% parents) recruited through a large respiratory medicine clinic in northern England, supported by the CT Trust. Represents n=15 invited for interview, reason for non-participation not given. *Unclear discrepancy within the study publication. A table and all results report the views and characteristics of n=10 participants; the methods report successful recruitment of only n=9.	No detail given. Main results report views on 'selective reproduction'. This is not explicitly defined, but the results discussion states '…support of a genetic carrier screening program being introduced for the condition they live with, whether this be a preconception genetic screening program or a prenatal screening program'	View on 'selective reproduction' of n=10 respondents: Approves: 30% Disapproves: 50% Conflicting: 20%

The 10 people with CF interviewed in this study have conflicting views on selective reproduction at a population level, with half disapproving.

Key themes within the categories of support or non-support were derived from the analysis and presented for participants with different genetic conditions. Only the 3 key themes explicitly identified from CF participants are presented below. The publication discusses anecdotes from 3 of the 10 respondents, one was supportive of selective reproduction and 2 were not. A 32-year-old mother gave her support in relation to 'the physical impact of the disease,' expressing the view that only people with CF, not doctors, can understand what it is really like to have CF and that 'we're

the only ones who can make these [reproductive screening] decisions accurately.'

Of the 2 respondents not in favour of selective reproduction, the views of one 32-year-old man related to 'valuing life affected by genetic disease.' He said that CF placed limitations on his life, but he couldn't say it affected his ability to take advantage of life. He saw a contradiction between 'championing and affirming the lives of people with CF' and a screening programme that 'opens the door to someone aborting me, or someone else with CF'. The non-support of another 58-year-old man related to 'the identity politics of genetic disease.' He had a child via a donor and was glad that his child was not a carrier and so 'wouldn't have to worry when it comes to her turn to have children.' However, while 'in an ideal world' he said he would be glad to see a decline in the disease through medical advances, he felt 'it's just a disability at the end of the day it doesn't dictate how your whole life's going to be.' His opposition to selective reproduction therefore centred on 'valuing the fetus *only* as a CF fetus' rather than for the potential value that future child could bring to life.

The study provides only a limited perspective of the views of 10 people with CF from a single UK clinic. It is not clear why 5 people chose not to participate in the interviews, for example, whether they may have been in favour, against or undecided about population-based screening. Overall it is very unclear how well these views may reflect those of the wider population with CF. Furthermore the participants were not specifically asked whether their views relate to antenatal or preconception screening. People with CF may have differing views about the two approaches. Overall this study cannot provide conclusive understanding about the views of people with CF towards population-wide antenatal screening.

On this basis the study would be unlikely to change the conclusions of this evidence review and Criterion 12 would remain 'not met.'

Review summary

Conclusions and implications for policy

The evidence to support a population-based antenatal screening programme for CF is not currently available. As such, the findings do not indicate that a change to the current policy should be made and antenatal screening for CF should not be recommended

The review did not identify the evidence needed to answer the key questions leaving several remaining uncertainties:

1. Data is available from the UK CF registry which shows that in 2016 CF affected 1 in 6276 people in the UK or 1.59 per 10,000 of the population. The incidence in 2016 was 1 in 3137 live births or 3.19 per 10,000 births per year. Birth incidence increased in 2007, the timing of introduction of NBS for CF. Since 2007 there has been no clear change in incidence, but prevalence has continued to rise since the Millennium. This suggests that survival may be improving. There has been little change in genotype prevalence over the years. Variant F508del is by far the most common variant carried by 90% of people with CF. Around 50% carry 2 copies of this variant (homozygotes). The UK CF registry covers 99% of people with CF seen in clinics across the UK, so is likely to give a true reflection of prevalence and incidence. Therefore, Criterion 1 on the prevalence and incidence of CF is met. There was no data on prevalence or incidence by ethnicity or on CF carrier prevalence.
2. There is consistent evidence from 15 large studies that gene variant class 1 to 5 is linked with phenotype in CF. F508del homozygotes and other people who carry 2 copies of a class 1-3 variant are likely to have pancreatic insufficiency and poor survival outlook. People who carry at least one class 4 or 5 variant are likely to have milder disease course with lower rates of pancreatic insufficiency and longer survival. There were similar but less consistent associations with lung function and age at diagnosis. However, across studies phenotype was highly variable for people with the same genotype or with variants in the same functional class. One study looked at the ability of genotype class to predict age at death. It found that while most people who die before age 30 years carry 2 severe class 1-3 variants, a third with these genotypes live beyond this age. Similarly around a third of people with at least one mild class 4 or 5 variants still die before 30. Studies generally found that around 90-100% of people with 2 class 1-3 variants including F508del homozygotes had pancreatic insufficiency, and were usually diagnosed prior to the age of 2 years. However, between 25% and 75% of people with at least one class 4 or 5 variant also had pancreatic insufficiency, and although diagnosis was usually later, it varied from childhood to adulthood. Therefore it

would not be possible to accurately predict individual disease course with any certainty based on genotype alone.

There are also several limitations to the evidence. Most studies are based on registry data and genotype or classification information was not available for typically half of the registry population. Therefore results may not represent the CF population as a whole. Few studies adjusted for treatment or care received and other confounding variables, increasing the risk of bias. Furthermore most cohorts date from over 20 years ago and may not be applicable today because treatment advances may have considerably altered prognosis. Finally, only a few potentially disease-causing CF variants have been widely studied, classified or included in prior antenatal screening panels. The phenotypic effects of many rare variants are unknown.

Overall, there is evidence of an association between genotype and phenotype. However, due to the variability in outcomes for individuals, risk of bias across studies (particularly relating to lack of genotyping and confounding), limited applicability to care today, and uncertain effects of rare variants, there is insufficient evidence to reliably predict the genotype-phenotype association. This degree of uncertainty is considered a reasonable price to pay in newborn screening as more babies will benefit than be harmed from screening. However, in antenatal testing where the option is to continue or terminate the pregnancy, a much higher degree of certainty is needed. Therefore this part of Criterion 1 on genotype-phenotype association is not met. Furthermore, in the newborn programme, mutation analysis is a second step only carried out for infants with immunoreactive trypsinogen levels above the cut-off (on 2 assays).

3. No studies have been published investigating antenatal screening in the UK since 2000. Only a single screening pilot has been conducted in Victoria, Australia. This study screened 3200 individuals and detected 106 carriers with a carrier frequency of 1 in 30. Subsequent screening of their partners (sequential testing) identified 6 pregnant carrier couples, all of whom accepted diagnostic testing. The 2 positive pregnancies (positive predictive value 33%) were both terminated, which is consistent with findings of pre-2000 screening pilots. There was no follow-up of screen-negatives so further test accuracy data was not available. This study also had limited applicability to the UK as it was a pay-for service, included preconception screening and tested for variants prevalent in the local population (not all of which are common in the UK). Pre-2000 UK pilots had also differed in the variants they tested for and the background literature indicates that there is as yet no well-established variant panel that could be used in an antenatal screening test for CF in the UK. Therefore Criteria 4 and 8 were not met.
4. No studies have assessed views on universal antenatal CF screening among the UK population. A sample of non-UK literature identified by the search included views of people taking part in the post-2000 Australian screening pilot. This generally indicated a lack of understanding of about CF screening,

for example, believing if you received a negative test result you were definitely not a carrier of any CF disease-causing variants; high levels of anxiety about antenatal diagnosis among couples who screened positive; and grief and regret over termination decisions. An additional Belgian study questioning views of people affected by CF (majority Catholic) found concerns that it would detract resources from CF and increase termination rates. These studies do not represent all of the international literature on screening views and are culturally-specific so cannot be generalised to the UK. On the basis of no UK evidence Criterion 12 is not met.

Further research may help to address the uncertainties around each of these 4 key questions:

1. Information on the carrier prevalence of CF variants among the general UK population, overall and by ethnicity. Information on the prevalence and incidence of CF by ethnicity.
2. Improved understanding of the phenotypic effects of rarer CF variants, and of the influence that modifier genes (other than *CFTR*), complex alleles (more than one disease-causing variant on the same allele) and environmental factors may have on genotype-phenotype relationships
3. To establish a panel of variants that could be used in antenatal screening in the UK and to conduct further antenatal screening pilots in the UK that use these variants. Such studies would benefit from conducting longer term follow-up and surveillance of all screen-negatives to give an indication of clinical sensitivity, specificity, positive and negative predictive values of the test
4. Study of the whether a population-wide antenatal screening programme is acceptable in the UK, to the population in general, to carriers and to people affected by CF.

Limitations

The search strategy was built on a protocol developed *a priori* for each of the 4 key questions. Searching was limited to 3 literature databases (4 for question 2 on genotype-phenotype association) and did not include grey literature resources for questions 3 and 4. Studies only available in non-English language, editorials abstracts, conference reports or poster presentations were not included. The reviewers were also unable to contact study authors or review non-published material. The systematic review on genotype-phenotype association has not analysed the effect of complex alleles (more than one variant on the same allele) or the influence of environmental factors or genes other than *CFTR* that may mediate the genotype-phenotype association.

Appendix 1 — Search strategy

Rapid review questions

Electronic databases

The search strategy for the 3 rapid review questions included searches of the databases shown in Table . MEDLINE, MEDLINE In-Process, MEDLINE Daily, Epub Ahead of Print and Embase.

Table 14. 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	Ovid SP	13/04/18	1946 to search date
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	13/04/18	CDSR: to search date

Search Terms

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

- Disease area: Cystic fibrosis
- Key questions terms
- Geographic terms

Search terms for MEDLINE, MEDLINE In-Process, MEDLINE Daily, Epub Ahead of Print and Embase are shown in Tables 15-17.

Key question 1: What is the prevalence of a) cystic fibrosis and b) cystic fibrosis carrier status in the UK and has it changed over time? [2000 to present, UK]

Table 15. Search strategy for MEDLINE, MEDLINE In-Process, MEDLINE Daily, Epub Ahead of Print and Embase Key question 1

Term Group	#	Search terms	Results
Disease area	1	(('cystic fibrosis' OR 'cf') NEAR/3 (carrier* OR heterozygote)):ab,ti	951
Disease area	2	'heterozygote'/de AND 'cystic fibrosis'/exp	1164
Disease area	3	'cystic fibrosis':ab,ti OR 'cf':ab,ti	88561
Disease area	4	'cystic fibrosis'/exp	63689
Disease area	5	#1 OR #2 OR #3 OR #4	102288
Key question terms	6	epidemiolog*:ab,ti OR inciden*:ab,ti OR prevalen*:ab,ti	2102764
Key question terms	7	'epidemiology'/de OR 'prevalence'/de OR 'incidence'/de	1023806
Key question terms	8	#6 OR #7	2399102
Geographic terms	9	britain:ti,ab OR british:ti,ab OR wales:ti,ab OR scotland:ti,ab OR england:ti,ab OR 'united kingdom':ti,ab OR uk:ti,ab OR gb:ti,ab	420026
Geographic terms	10	'united kingdom'/exp	405168
Geographic terms	11	#9 OR #10	656691
	12	#5 AND #8 AND #11	574
	13	#12 AND 'human'/de AND [english]/lim AND [2000-2018]/py	419

Key question 3: What genotypes/mutations do commercially available tests for antenatal screening of cystic fibrosis in the UK detect and how accurate are they? [2000 to present, UK/Europe/select countries]

Table 16. Search strategy for MEDLINE, MEDLINE In-Process, MEDLINE Daily, Epub Ahead of Print and Embase Key question 3

Term Group	#	Search terms	Results
Disease area	1	(((antenatal OR prenatal OR pregnan*) NEAR/3 (screen* OR test* OR diagnos* OR amniocentesis OR 'chorionic villus sampl*' OR cvs)):ab,ti) AND ('cystic fibrosis':ab,ti OR 'cf':ab,ti OR ((('cystic fibrosis' OR 'cf') NEAR/3 (carrier* OR heterozygot* OR parent* OR couple*)):ab,ti))	1019
Disease area	2	('heterozygote'/de AND 'cystic fibrosis'/exp OR 'cystic fibrosis'/exp) AND ('prenatal screening'/exp OR 'genetic screening'/exp OR 'mass screening'/exp OR 'amniocentesis'/exp OR 'chorion villus sampling'/de)	3735
Disease area	3	#1 OR #2	4369
Key question terms	4	'predictive value':ab,ti OR sensitivity:ab,ti OR specificity:ab,ti OR 'diagnostic accuracy':ab,ti OR diagnos*:ab,ti OR ((false NEAR/3 positive*):ti,ab) OR ((false NEAR/3 negative*):ti,ab) OR (((screen* OR diagnos* OR test*) NEAR/5 accura*):ti,ab) OR (((screen* OR diagnos* OR test*) NEAR/5 performance*):ti,ab)	4042568
Key question terms	5	'predictive value'/exp OR 'sensitivity and specificity'/exp OR 'diagnostic accuracy'/exp OR 'diagnosis'/exp	6287750

Key question terms	6	#4 OR #5	8179847
	7	#3 AND #6	3247
	8	#7 AND 'human'/de AND [english]/lim AND [2000-2018]/py	2015
Geographic terms	9	eu:ti,ab OR ((europe* NEAR/3 union):ti,ab) OR ((europe* NEAR/3 community):ti,ab) OR austria:ti,ab OR belgium:ti,ab OR bulgaria:ti,ab OR croatia:ti,ab OR cyprus:ti,ab OR ((czech NEAR/3 republic):ti,ab) OR denmark:ti,ab OR estonia:ti,ab OR finland:ti,ab OR france:ti,ab OR germany:ti,ab OR greece:ti,ab OR hungary:ti,ab OR ireland:ti,ab OR italy:ti,ab OR latvia:ti,ab OR lithuania:ti,ab OR luxembourg:ti,ab OR malta:ti,ab OR netherland:ti,ab OR poland:ti,ab OR portugal:ti,ab OR romania:ti,ab OR slovakia:ti,ab OR slovenia:ti,ab OR spain:ti,ab OR sweden:ti,ab OR britain:ti,ab OR wales:ti,ab OR scotland:ti,ab OR england:ti,ab OR 'northern ireland':ti,ab OR uk:ti,ab OR gb:ti,ab	981169
Geographic terms	10	'european union':de OR austria:de OR belgium:de OR bulgaria:de OR croatia:de OR cyprus:de OR 'czech republic':de OR denmark:de OR estonia:de OR finland:de OR france:de OR germany:de OR greece:de OR hungary:de OR ireland:de OR italy:de OR latvia:de OR lithuania:de OR luxembourg:de OR malta:de OR netherlands:de OR poland:de OR portugal:de OR romania:de OR slovakia:de OR slovenia:de OR spain:de OR sweden:de OR 'united kingdom'/exp	1216625
Geographic terms	11	#9 OR #10	1679792
Geographic terms	12	'australia and new zealand'/exp OR 'canada'/exp OR 'united states'/exp	1499244
Geographic terms	13	australia*:ti,ab OR 'new zealand':ti,ab OR canad*:ti,ab OR 'united states':ti,ab OR 'us':ti,ab	1022695
Geographic terms	14	#12 OR #13	2135748
Geographic terms	15	#11 OR #14	3694504
Disease + geographic	16	#8 AND #15	739

Key question 4: Is an antenatal cystic fibrosis screening programme acceptable in the UK? [2000 to present, UK]

Table 17. Search strategy for MEDLINE, MEDLINE In-Process, MEDLINE Daily, Epub Ahead of Print and Embase Key question 4

Term Group	#	Search terms	Results
Disease area	1	(((antenatal OR prenatal OR pregnan*) NEAR/3 (screen* OR test* OR diagnos* OR amniocentesis OR 'chorionic villus sampl*')):ab,ti) AND ('cystic fibrosis':ab,ti OR 'cf':ab,ti OR ((('cystic fibrosis' OR 'cf') NEAR/3 (carrier* OR heterozygote)):ab,ti))	1017
Disease area	2	('heterozygote'/de OR 'cystic fibrosis'/exp) AND ('prenatal screening'/exp OR 'genetic screening'/exp OR 'mass screening'/exp OR 'amniocentesis'/exp OR 'chorion villus sampling'/de)	8546
Disease area	3	#1 OR #2	9178

Key question terms	4	accepta*:ab,ti OR attitude*:ti,ab OR ((factor* NEAR/5 (influenc* OR affect*)):ti,ab) OR uptake*:ab,ti OR 'reproductive choice':ti,ab OR ((pregnancy NEAR/3 terminat*):ti,ab) OR decision?mak*:ti,ab OR preference*:ti,ab OR choice*:ti,ab	1558636
Key question terms	5	'social acceptance'/exp OR 'induced abortion'/de	23076
Key question terms	6	#4 OR #5	1577638
Geographic terms	7	britain:ti,ab OR british:ti,ab OR wales:ti,ab OR scotland:ti,ab OR england:ti,ab OR 'united kingdom':ti,ab OR uk:ti,ab OR gb:ti,ab	420026
Geographic terms	8	'united kingdom'/exp	405168
Geographic terms	9	#7 OR #8	656691
Geographic terms	10	#3 AND #6 AND #9	69
	11	#10 AND 'human'/de AND [english]/lim AND [2000-2018]/py	43

Results were imported into EndNote and de-duplicated.

Systematic review on genotype-phenotype association

Electronic databases

The search strategy for the systematic review included searches of the databases shown in Table .

Table 18. 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.com	11/05/18	1946 to search date
Embase	Embase.com	11/05/18	1974 to search date
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	11/05/18	To search date
Scopus	Scopus.com	11/05/18	1970 to search date

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: Cystic fibrosis
- Key questions terms: What are the genotype-phenotype associations in people with cystic fibrosis, including their clinical prognosis?

Search terms for MEDLINE, MEDLINE In-Process, MEDLINE Daily, Epub Ahead of Print and Embase via Embase.com are shown in Table 1, and search terms for Scopus are shown in Table 20.

Table 19. Search strategy for MEDLINE, MEDLINE In-Process, MEDLINE Daily, Epub Ahead of Print and Embase

Term Group	#	Search terms	Results
Disease area	1	'cystic fibrosis'/exp OR 'cystic fibrosis':ab,ti OR 'cf':ab,ti	102878
Key question terms	2	'genotype'/exp OR 'phenotype'/exp OR 'gene mutation'/exp	1294342
Key question terms	3	genotyp*:ti,ab OR phenotyp*:ti,ab OR mutation*:ti,ab OR 'genetic determin':ti,ab OR 'genetic risk factor*':ti,ab OR ((congenital* NEAR/3 absen*):ti,ab) OR cbavd:ti,ab OR regression:ti,ab	2273684
Key question terms	4	'association of congenital defects'/exp OR 'disease course'/exp OR 'prediction and forecasting'/exp OR 'mortality'/exp OR 'survival'/exp	3854014
Key question terms	5	((clinical NEAR/3 (manifestation* OR characteristic*)):ti,ab) OR correlation*:ti,ab OR relation*:ti,ab OR associat*:ti,ab OR predict*:ti,ab OR class*:ti,ab OR course:ti,ab OR declin*:ti,ab OR mortality:ti,ab OR ((disease* NEAR/3 outcome*):ti,ab) OR surviv*:ti,ab OR prognos*:ti,ab OR severity:ti,ab OR deteriorat*:ti,ab	10906221
Key question terms	6	'lung function':ti,ab OR 'lung infection*':ti,ab OR 'bowel obstruction*':ti,ab OR 'pancrea* insufficien*':ti,ab OR 'pancrea* sufficien*':ti,ab OR 'pancreatitis':ti,ab OR 'pulmonary function':ti,ab OR 'pulmonary disease':ti,ab OR ((respiratory NEAR/2 impair*):ti,ab) OR 'respiratory insufficien*':ti,ab OR 'distal intestinal obstruction':ti,ab OR 'forced expiratory volume':ti,ab OR 'forced expiration volume':ti,ab OR fev:ti,ab OR malnutrition:ti,ab OR malnourished:ti,ab OR 'nutritional status':ti,ab OR infertili*:ti,ab OR 'life expectancy':ti,ab OR 'quality of life':ti,ab OR 'respiratory function':ti,ab OR fev1:ti,ab OR hospitali?ation*:ti,ab OR absenteeism:ti,ab OR ((absence* NEAR/3 (school OR work)):ti,ab) OR (('time off' NEAR/3 (school OR work)):ti,ab) OR qol:ti,ab OR fvc:ti,ab OR 'forced vital capacity':ti,ab OR ((treatment NEAR/3 response):ti,ab)	1083545
Key question terms	7	'lung function'/exp OR 'lung infection'/exp OR 'pancreatic	1674323

		insufficiency'/exp OR 'pancreatitis'/exp OR 'intestine obstruction'/exp OR 'forced expiratory volume'/exp OR 'malnutrition'/exp OR 'infertility'/exp OR 'quality of life'/exp OR 'respiratory function'/exp OR 'hospitalization'/exp OR 'absenteeism'/exp OR 'school attendance'/exp	
Disease area and key question terms	8	#1 AND (#2 OR #3) AND (#4 OR #5 OR #6 OR #7)	14079
Study type and language limitations	9	#8 NOT ('editorial'/exp OR 'erratum'/exp OR 'letter'/exp OR 'note'/exp OR 'conference paper'/exp OR 'chapter'/it OR 'conference abstract'/it OR 'conference review'/it OR 'letter'/it OR 'note'/it) AND [english]/lim	7588

Table 20. Search strategy for Scopus

Term Group	#	Search terms	Results
Disease area, key question terms and study type/language limitations		(TITLE-ABS ((cystic fibrosis) OR (cf))) AND (TITLE-ABS (genotyp* OR phenotyp* OR mutation* OR (genetic determin*) OR (genetic risk factor*) OR (congenital* W/3 absen*) OR cbavd OR regression)) AND ((TITLE-ABS ((clinical W/3 (manifestation* OR characteristic*)) OR correlation* OR relation* OR associat* OR predict* OR class* OR course OR declin* OR mortality OR (disease* W/3 outcome*) OR surviv* OR prognos* OR severity OR deteriorat*)) OR (TITLE-ABS ((lung function*) OR (lung infection*) OR (bowel obstruction*) OR (pancrea* insufficien*) OR (pancrea* sufficien*) OR pancreatitis OR (pulmonary function*) OR (pulmonary disease) OR (respiratory W/2 impair*) OR (respiratory insufficien*) OR (distal intestinal obstruction) OR (forced expiratory volume) OR (forced expiration volume) OR (fev) OR malnutrition OR malnourished OR (nutritional status) OR infertili* OR (life expectancy) OR (quality of life) OR (respiratory function) OR (fev1) OR hospitalisation* OR absenteeism OR (absence* W/3 school) OR (absence* W/3 work) OR ((time off) W/3 school) OR ((time off) W/3 work) OR (qol) OR (fvc) OR (forced vital capacity) OR (treatment W/3 response) OR hospitali?ation))) AND (LIMIT-TO (DOCTYPE , "ar ") OR LIMIT-TO (DOCTYPE , " re ") OR LIMIT-TO (DOCTYPE , " ip "))	6,672

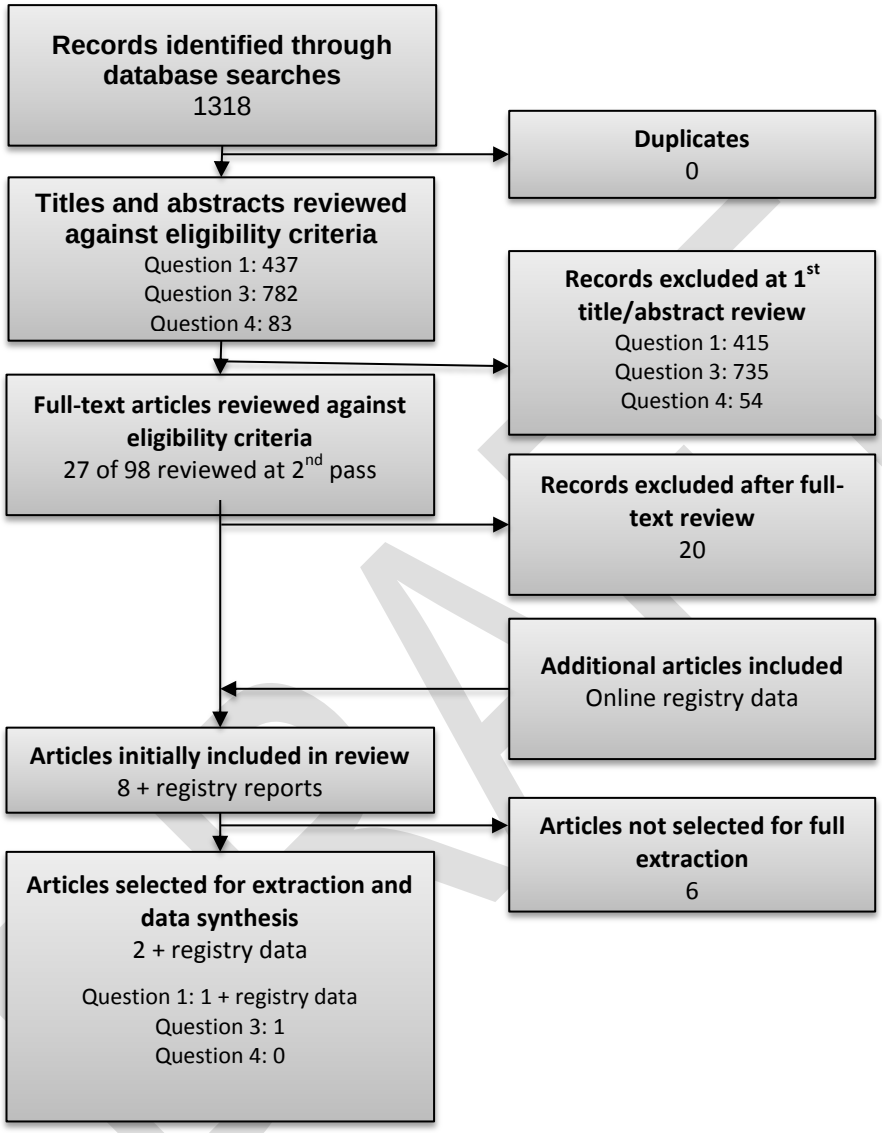
Results were imported into EndNote and de-duplicated.

Appendix 2 — Included and excluded studies

3 and 4 summarise the volume of publications included and excluded at each stage for, respectively, the rapid review and systematic review questions. Publications that were included or excluded after the review of full-text articles are detailed below.

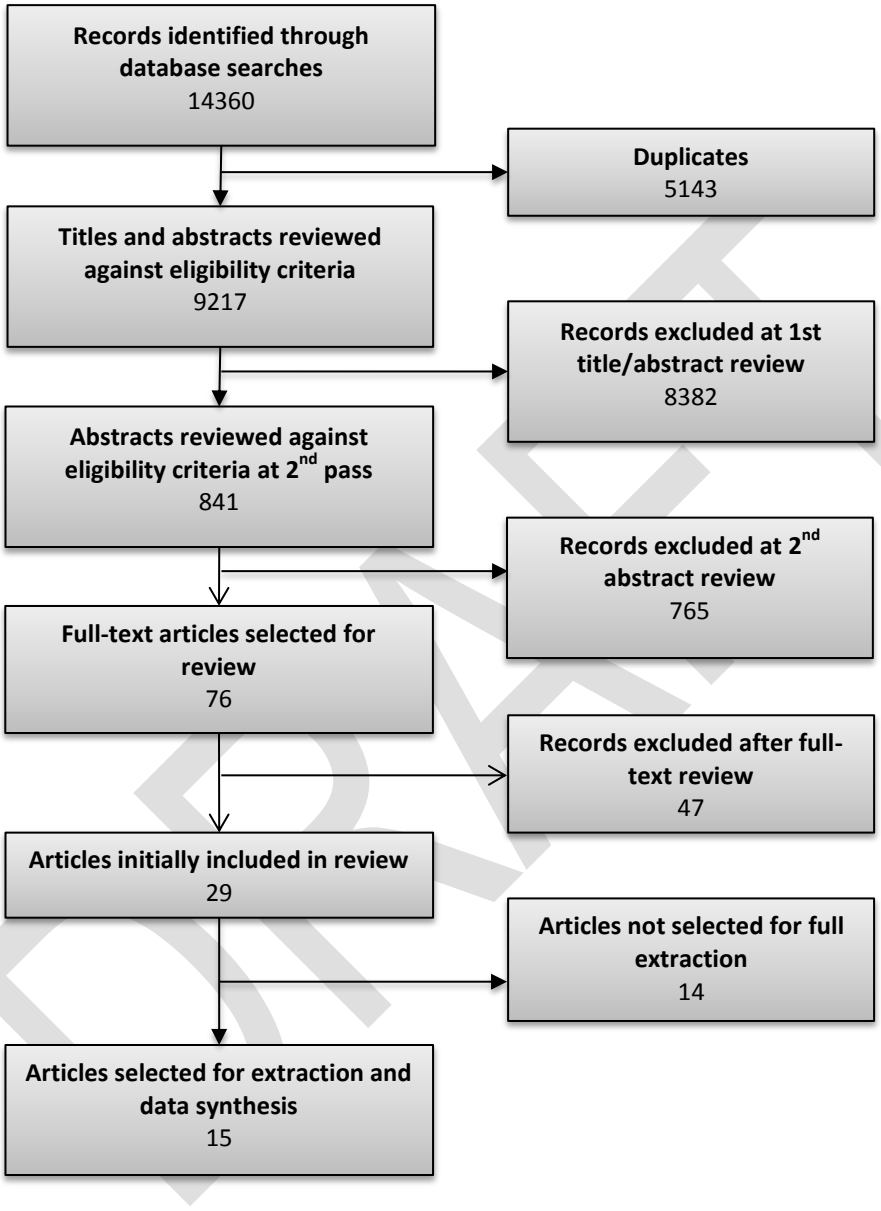
PRISMA flowchart for the rapid review questions

Figure 3. Summary of publications included and excluded at each stage of the rapid review



PRISMA flowchart for the systematic review question

Figure 4. Summary of publications included and excluded at each stage of the systematic review



Publications included after review of full-text articles

The publications and resources included after review of full-texts are summarised in Table 21. 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:

- 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 as listed in Table 2.
- For the prevalence and screening acceptability questions, only UK studies were relevant.
- For test question, studies would be prioritised if they considered a UK population, followed by studies from Western populations analogous to the UK. Only one non-UK pilot was identified. Contextual information from narrative reviews discussing potential screening panels relevant to the UK was discussed but does not provide evidence for the question.
- As no UK studies were identified for the acceptability question the themes for non-UK studies were summarised but do not provide evidence for the question.
- No restrictions were placed on study location or date for the systematic review question on genotype-phenotype association.

Publications not selected for extraction and data synthesis are clearly detailed in Table 21. Summary o.

Table 21. 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
UK CF Registry Annual reports	Q1	-	-	-	-	-
Hoo et al 2014[13]	Q1	-	-	-	-	-
Mckone et al 2006	Q2	-	-	-	-	
Mckone et al	Q2	-	-	-	-	

Study	The condition	The test	The intervention	The screening programme	Implementation criteria	Comments
2003						
Lai et al 2004	Q2	-	-	-	-	
O' Connor et al 2002	Q2	-	-	-	-	
Simmonds et al 2009	Q2	-	-	-	-	
Badet et al 2004	Q2	-	-	-	-	
Koch et al 2001	Q2	-	-	-	-	
Dewulf et al 2015	Q2	-	-	-	-	
Green et al 2010	Q2	-	-	-	-	
Radtke et al 2017	Q2	-	-	-	-	
The CF Consortium 1993	Q2	-	-	-	-	
Szczesniak et al 2017	Q2	-	-	-	-	
De Boeck and Zolin 2017	Q2	-	-	-	-	
Dugueperoux and De Braekeleer 2005	Q2	-	-	-	-	
MacKenzie et al 2017	Q2	-	-	-	-	
Massie et al 2009[42]	-	Q3	-	Q3	-	
Wald et al 2003[4]	-	Q3	-	Q3	-	Context only
Brennan et al 2016[5]	-	Q3	-	Q3	-	Context only
Ioannou et al 2010[47]	Q4	-	-	Q4	-	Non-UK limited applicability
Ioannou et al	Q4	-	-	Q4	-	Non-UK limited

Study	The condition	The test	The intervention	The screening programme	Implementation criteria	Comments
2015[48]						applicability
Janssens et al 2016[49]	Q4	-	-	Q4	-	Non-UK limited applicability
Maxwell et al 2014	Q4	-	-	Q4	-	Non-UK limited applicability

Publications excluded after review of full-text articles

Of the 103 publications selected for full text review, 67 were ultimately judged not to be relevant to this review. These publications, along with reasons for exclusion, are listed in Table 22.

Table 22. Publications excluded after review of full-text articles

Reference	Reason for exclusion
Q1 – prevalence	
Bosch B, Bilton D, Sosnay P, et al. Asian patients with CF: Does ethnicity influence our diagnostic criteria? Journal of Cystic Fibrosis. 2015;14:S42.	Abstract only.
De Boeck K, Zolin A, Cuppens H, et al. The relative frequency of CFTR mutation classes in European patients with cystic fibrosis. Journal of Cystic Fibrosis. 2014;13(4):403-9.	2009 data of cases reported to the European Registry for UK along with other countries. Contains figures oon the number with variants identified but information on frequency by class is by bar chart with no data.
Burgel PR, Bellis G, Olesen HV, et al. Future trends in cystic fibrosis demography in 34 European countries. European Respiratory Journal. 2015;46(1):133-41.	Modelling of European population estimates for 2025 based on 2009 data – different prevalence from UK registry reports and considered less relevant as data now goes to 2016.
Dodge JA, Lewis PA, Stanton M, et al. Cystic fibrosis mortality and survival in the UK: 1947-2003. European Respiratory Journal. 2007;29(3):522-6.	CF population by age and gender in 2003; also gives survival age of 3 yearly cohorts 1968-94 up to 2003. UK CF registry reports contain more information.
Edenborough FP, Morton AM. Cystic fibrosis - A guide for clinicians in reproductive and obstetric medicine. Fetal and Maternal Medicine Review. 2010;21(1):36-54.	Not possible to access publication, though expected to give background only.
Farrell PM. The prevalence of cystic fibrosis in the European Union. Journal of Cystic Fibrosis. 2008;7(5):450-3.	ECFSPR data from 2004 alongside literature review to 2006: UK data cited to Dodge: UK CF registry reports contain more information.
Goss CH. Country to country variation: What can be learnt from national cystic fibrosis registries. Current Opinion in Pulmonary Medicine. 2015;21(6):585-90.	Narrative review with no methods. Cites only UK CF registry for UK population homo or heterozygous for F508del: variant data contained in registry reports.

Hoo ZH, Wildman M, Teare MD. CF registry mortality analysis to understand the effects of widespread genetic testing on the trend of median age at death: Is the increased life-expectancy related to increased prevalence of mild phenotypes? Journal of Cystic Fibrosis. 2013;12:S140.	Abstract only.
Ioannou L, McClaren BJ, Massie J, et al. Population-based carrier screening for cystic fibrosis: A systematic review of 23 years of research. Genetics in Medicine. 2014;16(3):207-16.	Doesn't review prevalence data, only provides general background information on European population estimates.
Palomaki GE, FitzSimmons SC, Haddow JE. Clinical sensitivity of prenatal screening for cystic fibrosis via CFTR carrier testing in a United States panethnic population. Genetics in Medicine. 2004;6(5):405-14.	Analyses US studies reporting variant frequency of 25 ACMG panel among people of different ethnicities. No data relevant to the UK.
Q2 – genotype-phenotype association	
Al-Jader LN, Meredith AL, Ryley HC, Cheadle JP, Maguire S, Owen G, et al. Severity of chest disease in cystic fibrosis patients in relation to their genotypes. Journal of Medical Genetics. 1992;29(12):883-7.	Welsh centre analysis of decline in lung function by age for patients homozygous or heterozygous for F508del but only 76 patients with data.
Alvarez AE, Ribeiro AF, Hesselm G, Bertuzzo CS, Ribeiro JD. Cystic fibrosis at a Brazilian center of excellence: Clinical and laboratory characteristics of 104 patients and their association with genotype and disease severity. Jornal de Pediatria. 2004;80(5):371-9.	Brazilian centre. Only gives the statistical correlation for presence of F508del in homozygous or heterozygous forms for different variables. Only 78 patients in analysis and doesn't give any values for homozygotes or heterozygotes.
Augarten A, Tov AB, Madgar I, Barak A, Akons H, Laufer J, et al. The changing face of the exocrine pancreas in cystic fibrosis: The correlation between pancreatic status, pancreatitis and cystic fibrosis genotype. European Journal of Gastroenterology and Hepatology. 2008;20(3):164-8.	Israel national centres 505 patients, 128 of whom were pancreatic sufficient. Lists the genotypes of those with pancreatic sufficiency, but has carried out no analysis.
Bonizzato A, Bisceglia L, Marigo C, Nicolis E, Bombieri C, Castellani C, et al. Analysis of the complete coding region of the CFTR gene in a cohort of CF patients from North-Eastern Italy: Identification of 90% of the mutations. Human Genetics. 1995;95(4):397-402.	No statistical analysis. Just lists FEV1 by genotype for 59 patients at one Italian centre.
Borgo G, Mastella G, Gasparini P, Zorzanello A, Doro R, Pignatti PF. Pancreatic function and gene deletion F508 in cystic fibrosis. Journal of	Italian centre 123 patients. Statistical analysis is only for frequency of F508del among chromosomes of those

Medical Genetics. 1990;27(11):665-9.	pancreatic sufficient/insufficient, rather than informing whether the genotype was homozygous or heterozygous.
Cawood T. Cystic fibrosis-related diabetes in adults. Irish Medical Journal. 2006;99(3).	150 patients attending Irish centre, 81 with diabetes remainder without. Lists proportions carrying F508del, R117H, G551D and 'other' among these samples. Small size and little meaningful interpretation can be drawn to inform screening decisions.
Cipolli M, Castellani C, Wilcken B, Massie J, McKay K, Gruca M, et al. Pancreatic phenotype in infants with cystic fibrosis identified by mutation screening. Archives of Disease in Childhood. 2007;92(10):842-6.	Newborn screening samples from Italian and Austrlain centres: 315. Gives proportions pancreatic sufficient/insufficient by groups of F508del homozygotes, F508del compound heterozygotes and non- F508del compound heterozygotes – each by whether the other variant was mild/severe or unknown. Small numbers in each group and no statistical analysis.
Comer DM, Ennis M, McDowell C, Beattie D, Rendall J, Hall V, et al. Clinical phenotype of cystic fibrosis patients with the G551D mutation. QJM. 2009;102(11):793-8.	Belfast centre 101 patients grouped F508del homozygotes vs G551D compound vs R117H compound vs G551D/R117H. Primarily excluded as small study and clear how representative these people are (e.g. as appose to registry study that has identified all people with these variants).
Cotellessa M, Minicucci L, Diana MC, Prigione F, Di Febbraro L, Gagliardini R, et al. Phenotype/genotype correlation and cystic fibrosis related diabetes mellitus (Italian Multicenter Study). Journal of Pediatric Endocrinology and Metabolism. 2000;13(8):1087-93.	Italian multicentre study of 1229: 141 with diabetes. Gives the proportion of those with and without diabetes carrying 5 different variants, and of those with and without classical diabetes presentation. One of the variants shows significantly higher prevalence from control population (5 v 2%). Given the small number it's hard to know how reliable the analysis is. Also unclear whether other variants than the 5 selected for testing here could also carry diabetes risk.
De Bie I, Agatep R, Scott P, Ruchon A. Report on the p.Ser489X (p.Ser489) CFTR mutation, a variant with severe associated phenotype and high prevalence in a Quebec French-Canadian cystic fibrosis patient	Reporting symptoms for only 13 people carrying the Ser489X variant. Carried by 0.1% of UK CF population, but hasn't been included in any variant panels.

population. Genetics in Medicine. 2012;14(10):883-6.	
Dawson KP, Frossard PM, Al-Awar B. Disease severity associated with cystic fibrosis mutations deltaF508 and S549R(T-->G). Eastern Mediterranean health journal = La revue de santé de la Méditerranée orientale = al-Majallah al-ihhīyah li-sharq al-mutawassi. 2001;7(6):975-80.	Only 25 patients in the study, unclear from abstract.
De Arce M, O'Brien S, Hegarty J, O'Mahoney SM, Cashman SM, Martinez A, et al. Deletion ΔF508 and clinical expression of cystic fibrosis-related liver disease. Clinical Genetics. 1992;42(5):271-2.	Brief paper from Irish centre of 108 patients: 20 with liver disease. Analyses the proportion homozygous or heterozygous for F508del and finds no significant difference. Small size indicates this may not reliably exclude genotype association.
De Boeck K, Weren M, Proesmans M, Kerem E. Pancreatitis among patients with cystic fibrosis: Correlation with pancreatic status and genotype. Pediatrics. 2005;115(4):e463-e9.	Doctors asked to provide data through CF Thematic Network or European CF foundation on patients with pancreatitis. N=3306 total and n=61 with pancreatitis. Gives a long list of genotypes for these people by status sufficient, insufficient, PI after PS and unknown status. No analysis and minimal could be interpreted. Also starting status by pancreatitis rather than assessing gene association with sufficiency.
Duguépéroux I, De Braekeleer M. Genotype-phenotype relationship for five CFTR mutations frequently identified in western France. Journal of cystic fibrosis. 2004;3(4):259-63.	Analysis of French registry for patients carrying one of 5 variants of which delI507 (n=22) and 1078/delT (n=23) have been included in screening panels to date. However, the study involved comparison to matched groups of F508del homozygotes and their phenotypic values are not given.
Dupuis A, Keenan K, Ooi CY, Dorfman R, Sontag MK, Naehrlich L, et al. Prevalence of meconium ileus marks the severity of mutations of the Cystic Fibrosis Transmembrane Conductance Regulator (CFTR) gene. Genetics in Medicine. 2016;18(4):333-40.	Canadian Gene Consortium: includes 2,492 Canadian patients looking at the prevalence of meconium ileus for F508del heterozygotes. Data is available as a long list of scores according to variant combination. Excluded primarily as meconium ileus wasn't selected as a clinical outcome to inform likely prognosis in the scenario of antenatal screening.
Durno C, Corey M, Zielenski J, Tullis E, Tsui LC, Durie P. Genotype and phenotype correlations in patients with cystic fibrosis and pancreatitis. Gastroenterology. 2002;123(6):1857-64.	Canada CF database 110 with pancreatic sufficiency, 19 with pancreatitis. Lists genotypes of those with sufficiency with and without pancreatitis but has no

	statistical analysis.
Duthie A, Doherty DG, Williams C, Scott-Jupp R, Warner JO, Tanner MS, et al. Genotype analysis for ΔF508, G551D and R553X mutations in children and young adults with cystic fibrosis with and without chronic liver disease. Hepatology. 1992;15(4):660-4.	111 from UK hospitals, 63 no liver disease, 29 portal hypertension, 19 liver disease. Has tested only for F508del, G551D and R553X variants. Gives proportions in the 3 groups with different genotypes: F508del homozygotes, compound with G551D and R553X, then several others compound with unidentified variants. Excluded on size and limited relevance for informing screening decisions.
Feingold J, Guilloud-Bataille M, Albertini, et al. Genetic comparisons of patients with cystic fibrosis with or without meconium ileus. Annales de Genetique. 1999;42(3):147-50	Unable to access full text.
Gilljam M, Ellis L, Corey M, Zielenski J, Durie P, Tullis DE. Clinical manifestations of cystic fibrosis among patients with diagnosis in adulthood. Chest. 2004;126(4):1215-24.	Only lists genotypes for 46 patients diagnosed as adults. No statistical comparison to younger patients and considered to have minimal value for informing potential screening programmes.
Gorinova YV, Savostyanov KV, Pushkov AA, Nikitin AG, Pen'Kov EL, Krasovskiy SA, et al. Genotype-phenotype correlations of the course of cystic fibrosis in Russian children. the first description of eleven new mutations. Voprosy Sovremennoi Pediatrii - Current Pediatrics. 2018;17(1):61-9.	Only abstract and tables available in English language. Traits grouped according to class 1, missense or 'other' variants in table. Excluded as no further detail is available to analyse methods.
Hergersberg M, Balakrishnan J, Bettecken T, Chevalier-Porst F, Brägger C, Burger R, et al. A new mutation, 3905insT, accounts for 4.8% of 1173 CF chromosomes in Switzerland and causes a severe phenotype. Human Genetics. 1997;100(2):220-3.	Variant analysis of 606 CF families in Switzerland. Reports variants and frequency. 56 chromosomes found to have 3905insT. Simply states that these patients had pancreatic insufficiency and focuses on one female patient with the variant. Variant is carried by <0.1% of people with CF in the UK (<5 individuals in 2016 registry).
Hilman BC. Genetic and immunologic aspects of cystic fibrosis. Annals of Allergy, Asthma and Immunology. 1997;79(5):379-94.	Narrative review only, unclear from abstract.
Hoo ZH, Wildman MJ, Teare MD. Exploration of the impact of 'mild phenotypes' on median age at death in the UK CF registry. Respiratory Medicine. 2014;108(5):716-21.	No analysis of genotype, looking at pancreatic sufficiency and association with survival only.
Hubert D, Bienvenu T, Desmazes-Dufeu N, Fajac I, Lacronique J, Matran R, et al. Genotype-phenotype relationships in a cohort of adult cystic	France centre 110 patients. Small study with four groupings with mix of genotypes and not like classic

fibrosis patients. European Respiratory Journal. 1996;9(11):2207-14.	groupings. Difficult to interpret for the purpose of this review.
Keller BM, Casaulta Aebischer C, Kraemer R, Schöni MH. Growth in prepubertal children with cystic fibrosis, homozygous for the ΔF508 mutation. Journal of Cystic Fibrosis. 2003;2(2):76-83.	Unclear from abstract but including only 35 children and no analysis by genotype.
Kraemer R, Baldwin DN, Ammann RA, Frey U, Gallati S. Progression of pulmonary hyperinflation and trapped gas associated with genetic and environmental factors in children with cystic fibrosis. Respiratory Research. 2006;7.	USCFPR small sample of 152 with lung function data over 28 years since 1978 divided into 4 groups which are not compatible for analysis alongside other studies, e.g. trapped gas, lung clearance index, pulmonary hyperinflation.
Kubesch P, Dork T, Wulbrand U, Kalin N, Neumann T, Wulf B, et al. Genetic determinants of airways' colonisation with Pseudomonas aeruginosa in cystic fibrosis. Lancet. 1993;341(8839):189-93.	German centre 267 patients put into the researchers' groupings which have difficult interpretation for the purpose of this review.
Lanng S, Schwartz M, Thorsteinsson B, Koch C. Endocrine and exocrine pancreatic function and the ΔF508 mutation in cystic fibrosis. Clinical Genetics. 1991;40(5):345-8.	Danish single centre study of 215 patients (including 15 sibling pairs). Compares F508del homozygotes and heterozygotes (any) for number of Pancaps enzyme tablets per day per age band. Small study with little compatibility with other pancreatic sufficiency assessments.
Lewis C, Blackman SM, Nelson A, Oberdorfer E, Wells D, Dunitz J, et al. Diabetes-related mortality in adults with cystic fibrosis: Role of genotype and sex. American Journal of Respiratory and Critical Care Medicine. 2015;191(2):194-200.	US centre case control 462: half with diabetes. Analysing diabetes-related mortality rates for those with and without mild or severe class variant. Excluded as not a standard analysis between genotype and survival, but with diabetes as a moderating factor. Therefore minimal value for informing screening decisions.
Lucarelli M, Bruno SM, Pierandrei S, Ferraguti G, Testino G, Truglio G, et al. The Impact on Genetic Testing of Mutational Patterns of CFTR Gene in Different Clinical Macrocategories of Cystic Fibrosis. Journal of Molecular Diagnostics. 2016;18(4):554-65.	Looking at assay detection rate. No genotype-phenotype analysis.
Maisonneuve P, Campbell IP, Durie P, Lowenfels AB. Pancreatitis in hispanic patients with cystic fibrosis carrying the R334W mutation. Clinical Gastroenterology and Hepatology. 2004;2(6):504-9.	Large US registry study of 17,871. Looking at the risk of pancreatitis by number of attacks in people with genotypes associated with pancreatic sufficiency/insufficiency. Demonstrates higher risk of pancreatitis among those with mild variants/sufficiency compared to severe variants/insufficiency. Also

	analyses frequency of variants among Hispanics compared with the general US population. Overall the contained information was considered to have limited relevance for informing screening decisions (presence of severe variant/insufficiency being the more pertinent prognostic factor for long-term outcomes).
Ooi CY, Dorfman R, Cipolli M, Gonska T, Castellani C, Keenan K, et al. Type of CFTR mutation determines risk of pancreatitis in patients with cystic fibrosis. Gastroenterology. 2011;140(1):153-61.	227 patients from Canadian Consortium for CF Genetic Studies and Verona CF Centre database who were pancreatic sufficient and documented as having pancreatitis/no pancreatitis. Previously established pancreatic insufficiency prevalence score were used as the surrogate for the severity of the patient's genotype (unclear whether single or both variants). Then analysed proportions with those genotypes among pancreatic sufficient patients with/without pancreatitis. Not a direct analysis of the association between genotypes and pancreatic sufficiency but exploring the higher risk of pancreatitis within these groups which was considered less relevant for informing potential screening programmes.
Osborne L, Santis G, Schwarz M, Klinger K, Dork T, McIntosh I, et al. Incidence and expression of the N1303K mutation of the cystic fibrosis (CFTR) gene. Human Genetics. 1992;89(6):653-8.	Collaborative international study identifying 216 chromosomes carrying of the N1303K variant. The variant has been included in screening panels but there appear some reliability issues. The study lists pancreatic status, FEV1 and sputum colonisation for people with each specific genotype. Could inform likelihood of phenotype for this variant but there is considerable missing data. Pancreatic status is only available for 97/206, nearly all of whom were insufficient. But unclear whether missing 50% could have sufficiency (rather than indicate all with this variant are insufficient). Then FEV1 is only available for 39/206.
Rosenecker J. Relations between the frequency of the DeltaF 508 mutation and the course of pulmonary disease in cystic fibrosis patients infected with Pseudomonas aeruginosa. European journal of medical research. 2000;5(8):356-9.	Unable to access full text.

Salvatore D, Buzzetti R, Baldo E, Forneris MP, Lucidi V, Manunza D, et al. An overview of international literature from cystic fibrosis registries. Part 3. Disease incidence, genotype/phenotype correlation, microbiology, pregnancy, clinical complications, lung transplantation, and miscellanea. Journal of Cystic Fibrosis. 2011;10(2):71-85.	Based on analysis of studies starting from national CF registries that have used data to describe aspects of the disease or advance research. Lists and gives brief discussion of 15 looking at "genetics". Useful to cross check against included studies, but there doesn't appear a specific set of inclusion criteria other than that the studies commenced from the registry and looked at genetics.
Santis G, Osborne L, Knight RA, et al. Independent genetic determinants of pancreatic and pulmonary status in cystic fibrosis. Lancet. 1990;336(8723):1081-4.	UK centre, 54 families with ≥2 siblings with CF. Genotyping performed for 105 people. Lists the genotypes among those pancreatic sufficient/insufficient. No statistical analysis. Aside from 51 F508del homozygotes few by other genotype. Limited conclusions can be drawn, also selective sample.
Schaedel C, De Monestrol I, Hjelte L, Johannesson M, Kornfält R, Lindblad A, et al. Predictors of deterioration of lung function in cystic fibrosis. Pediatric Pulmonology. 2002;33(6):483-91.	475 patients in Sweden put into the researchers' 4 groupings (F508del homozygotes vs severe/severe vs missense/severe or missense/missense vs unknown) which have difficult interpretation for the purpose of this review.
Selvadurai HC, McKay KO, Blimkie CJ, Cooper PJ, Mellis CM, Van Asperen PP. The relationship between genotype and exercise tolerance in children with cystic fibrosis. American Journal of Respiratory and Critical Care Medicine. 2002;165(6):762-5.	Only 97 child participants in study on exercise capacity. Pancreatic sufficiency, lung function and BMI analysed by class of second variant but excluded based on size.
Sims EJ, Green MW, Mehta A. Decreased lung function in female but not male subjects with established cystic fibrosis-related diabetes. Diabetes care. 2005;28(7):1581-7.	UKCF database 2000-02 large sample n=2640 but not straightforward genotype-phenotype analysis. Looks at how diabetic status predicts lung function by gender in all genotypes, those homozygous for F508del and age-matched.
Sosnay PR, Siklosi KR, Van Goor F, Kaniecki K, Yu H, Sharma N, et al. Defining the disease liability of variants in the cystic fibrosis transmembrane conductance regulator gene. Nature Genetics. 2013;45(10):1160-7.	CFTR2 project aiming to increase gene variants that have propensity to cause disease. Description of functional and penetrance analysis. No data on phenotype association with specific variants.
Tabori H, Arnold C, Jaudszus A, Mentzel HJ, Renz DM, Reinsch S, et al. Abdominal symptoms in cystic fibrosis and their relation to genotype,	131 patients looking at number of gastrointestinal symptoms. Reports rate of symptoms in those with mild

history, clinical and laboratory findings. PLoS ONE. 2017;12(5).	genotypes as informed by pancreatic insufficiency score rather than by genotypes.
Terlizzi V, Tosco A, Tomaiuolo R, et al. Prediction of acute pancreatitis risk based on PIP score in children with cystic fibrosis. Journal of cystic fibrosis. 2014;13(5):579-84.	Italian centre: 185 paediatric patients. Uses the same method to assign and compare pancreatic insufficiency prevalence scores as validated by Ooi et al. Lists the genotype and score of the 10 pancreatic sufficient patients who developed pancreatitis. Similar to Ooi, not a direct analysis of the association between genotypes and pancreatic sufficiency but exploring the higher risk of pancreatitis within these groups.
Tümmler B, Dörk T, Kubesch P, Fislage R, Kälin N, Neumann T, et al. Cystic fibrosis: the impact of analytical technology for genotype-phenotype studies. Clinica Chimica Acta. 1993;217(1):23-8.	Variant analysis in German and Turkish patients. General narrative discussion around the features linked with different genotype but no quantitative analysis.
Vandevanter DR, Pasta DJ, Konstan MW. Improvements in lung function and height among cohorts of 6-year-olds with cystic fibrosis from 1994 to 2012. Journal of Pediatrics. 2014;165(6):1091-7.e2.	US registry study (n=11,670) of successive cohorts of 6-year-olds. Regression of means for lung function for period 1994-2012. Shows no change in FEV1/FVC over the period, including when restricting the analysis to F508del homozygotes, but no analysis of genotype-phenotype association.
Van De Weert-van Leeuwen PB, Slieker MG, Hulzebos HJ, Kruitwagen CLJJ, Van Der Ent CK, Arets HGM. Chronic infection and inflammation affect exercise capacity in cystic fibrosis. European Respiratory Journal. 2012;39(4):893-8.	Netherlands centre 149 adolescents attending exercise assessments over 10 year period. Has analysed whether they carry at least one mild or two severe variants. However, analysis is multilinear mixed model assessing effect of CFTR variant class, chronic *P. aeruginosa* and inflammation on rate of decline of FEV1 over assessment period. No separate analysis of genotype.
Walkowiak J, Herzig KH, Witt M, Pogorzelski A, Piotrowski R, Barra E, et al. Analysis of exocrine pancreatic function in cystic fibrosis: One mild CFTR mutation does not exclude pancreatic insufficiency. European Journal of Clinical Investigation. 2001;31(9):796-801.	Poland: 394 patients seen 1993-2000 who had been genotyped. Looking at fecal elastase-1 concentration as a test for pancreatic sufficiency and lists these values by genotype. Little compatibility of outcome for analysis alongside other studies.
Zergollern L, Stavljenic A, Barisic I, Sertic J. The ΔF508 mutation and genotype-phenotype correlation in Croatian cystic fibrosis families.	Unable to access full text.

Periodicum Biologorum. 1993;95(3):359-61.	
Q3 – test accuracy	
Castellani C, Picci L, Tamanini A, et al. Association between carrier screening and incidence of cystic fibrosis. JAMA - Journal of the American Medical Association. 2009;302(23):2573-9.	Primarily looking at incidence in Italy and relationship to type of screening offered. Combines high risk and population screening with no data specific to population-based.
D'Apice MR, Novelli G, Sangiuolo F. Diagnostic CFTR mutation analysis. Expert Opinion on Medical Diagnostics. 2008;2(2):191-205.	Not able to access publication and would only be providing background on available tests.
Deeb KK, Metcalf JD, Sesock KM, et al. The c.1364C>A (p.A455E) mutation in the CFTR pseudogene results in an incorrectly assigned carrier status by a commonly used screening platform. Journal of Molecular Diagnostics. 2015;17(4):360-5.	Single case study reporting that short amplification-based carrier tests can lead to false positives.
Heim RA, Sugarman EA, Allitto BA. Improved detection of cystic fibrosis mutations in the heterogeneous U.S. population using an expanded, pan-ethnic mutation panel. Genetics in Medicine. 2001;3(3):168-76.	Variant analysis of 3000 US CF patients. Presenting the case for expanding the variant panel for population screening to 64 variants.
Kanavakis E, Efthymiadou A, Strofalis S, et al. Cystic fibrosis in Greece: Molecular diagnosis, haplotypes, prenatal diagnosis and carrier identification amongst high-risk individuals. Clinical Genetics. 2003;63(5):400-9.	Analyses variant frequency of 437 Greek CF patients. Separately reports variants identified in 116 antenatal screens, mostly tested on the basis of family history.
Le Maréchal C, Audrézet MP, Quéré I, et al. Complete and rapid scanning of the cystic fibrosis transmembrane conductance regulator (CFTR) gene by denaturing high-performance liquid chromatography (D-HPLC): Major implications for genetic counselling. Human Genetics. 2001;108(4):290-8.	Analysing accuracy of D-HPLC scanning technique. No relevant data.
Palomaki GE, FitzSimmons SC, Haddow JE. Clinical sensitivity of prenatal screening for cystic fibrosis via CFTR carrier testing in a United States panethnic population. Genetics in Medicine. 2004;6(5):405-14.	Analyses US studies reporting variant frequency of 25 ACMG panel among people of different ethnicities. No data relevant to the UK.
Trevisiol C, Boniotto M, Giglio L, et al. MBL2 polymorphisms screening in a regional Italian CF Center. Journal of Cystic Fibrosis. 2005;4(3):189-91.	Single centre cohort of patients with CF and their variants. No data relevant to question.

Q4 – acceptability	
Antoniadi T, Pampanos A, Petersen MB. Attitudes towards reproductive issues and career testing among adult patients and patents of children with cystic fibrosis (CF). Prenatal Diagnosis. 2001;21(1):1-9.	1997 Dutch National survey of CF patients or parents. Most relate their own antenatal decisions and reproductive choices. Single question asked about whether population screening should be offered (Y/N) to those *planning* pregnancy but not antenatal. Excluded on this basis, also limited relevance to current UK.
Beard CA, Amor DJ, Di Pietro L, et al. "I'm Healthy, It's Not Going To Be Me": Exploring experiences of carriers identified through a population reproductive genetic carrier screening panel in Australia. American Journal of Medical Genetics, Part A. 2016;170(8):2052-9.	Excluded on size. 3 women taking part in Australian programme interviewed on their feelings of being told they're carriers. 2 asked about views on population screen. Ioannou selected as contains data on all identified carriers.
Bruni T, Mameli M, Pravettoni G, et al. Cystic fibrosis carrier screening in Veneto (Italy): An ethical analysis. Medicine, Health Care and Philosophy. 2012;15(3):321-8.	Narrative of authors' views on the potential effects of antenatal and other forms of screening on CF incidence and reproductive decisions.
De Braekeleer M, Bellis G, Rault G, et al. Reproductive attitudes of couples having a child with cystic fibrosis in Saguenay-Lac-Saint-Jean (Quebec, Canada). Annales de Genetique. 2000;43(2):93-7.	Parents of CF children. Their personal views on subsequent use of antenatal diagnosis and reproductive decisions. No data relevant to views on population screening.
De Braekeleer M, Rault G, Bellis G. Reproductive attitudes of couples having a child with cystic fibrosis in Brittany (France). Journal of Human Genetics. 2004;49(6):285-9.	207 adults with CF or parents of CF children. Assessing effects on their own reproductive decisions and whether they'd use antenatal diagnosis. No data on views on population screening.
Henneman L, Bramsen I, Van Os TA et al: Attitudes towards reproductive issues and carrier testing among adult patients and parents of children with cystic fibrosis (CF). Prenat Diagn 2001; 21:1–9.	Only questions participants own reproductive decisions and their views on population screening for people *planning* pregnancy, so nothing relevant to antenatal.
Ioannou L, McClaren BJ, Massie J, et al. Population-based carrier screening for cystic fibrosis: A systematic review of 23 years of research. Genetics in Medicine. 2014;16(3):207-16.	Non-specific data on attitudes of members of general population about when screening should be offered (eg pre-pregnancy or pregnancy) and how they'd like to be given information (eg by leaflet or in-person).

Pisnoli L, O'Connor A, Goldsmith L, et al. Impact of fetal or child loss on parents' perceptions of non-invasive prenatal diagnosis for autosomal recessive conditions. Midwifery. 2016;34:105-10.	Interviews parents of CF children. Gives quotations on effects on them and their personal views on antenatal diagnosis and reproductive decisions. No data relevant to views on population screening.
Wright KF, Bryant LD, Morley S, et al. Presenting life with cystic fibrosis: a Q-methodological approach to developing balanced, experience-based prenatal screening information. Health expectations : an international journal of public participation in health care and health policy. 2015;18(5):1349-62.	Aspects of life that those affected by CF consider most important to include information on in antenatal screening (eg QoL effects). Nothing relevant to views on population screening or decisions.

Appendix 3 — Summary and appraisal of individual studies

Data Extraction

Criterion 1: question 1: prevalence

Table 23. Total registered cases per year in the UK CF registry (data from annual reports)

	2002	2003	2004	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Total registered cases	6932	6861	7046	8080	8513	9029	9385	9749	10078	10338	10583	10810	10461
UK population[14]	59365700	59636700	59950400	61319100	61823800	62260500	62759500	63285100	63705000	64105700	64596800	65110000	65648100
Prevalence per year	**1 in 8564**	**1 in 8692**	**1 in 8508**	**1 in 7589**	**1 in 7262**	**1 in 6896**	**1 in 6687**	**1 in 6491**	**1 in 6321**	**1 in 6201**	**1 in 6104**	**1 in 6023**	**1 in 6276**
Per 10,000 population	**1.17**	**1.15**	**1.18**	**1.32**	**1.38**	**1.45**	**1.50**	**1.54**	**1.58**	**1.61**	**1.64**	**1.66**	**1.59**
New diagnoses (NBS/other)	159	142	164	239	235	261	301	261	285	301	291	300	247
NBS diagnoses	-	-	-	-	-	-	189 (63%)	155 (59%)	213 (75%)	177 (59%)	164 (56%)	168 (56%)	180 (73%)
Total UK live births	668777	695549	715966	772245	794383	790204	807271	807776	812970	778803	776352	777167	774835
Incidence per year	**1 in 4206**	**1 in 4898**	**1 in 4366**	**1 in 3231**	**1 in 3380**	**1 in 3028**	**1 in 2682**	**1 in 3095**	**1 in 2853**	**1 in 2587**	**1 in 2668**	**1 in 2591**	**1 in 3137**
Per 10,000 population	**2.38**	**2.04**	**2.29**	**3.09**	**2.96**	**3.30**	**3.73**	**3.23**	**3.51**	**3.86**	**3.75**	**3.86**	**3.19**

- No data available for years 2000-01 and 2005-06.
- Total registered are people diagnosed with CF, seen within the past 2 years and alive at 1st January of that year.
- Numbers are reported to have decreased in 2016 due to data clearing. Registered patients who did not have data submitted in 2016 were followed up and those who were no longer being cared for within the NHS (for example, had moved abroad) were marked as 'inactive' and excluded.

- New diagnoses are based on data from the most recent reports as some diagnoses are added after data entry closure each year so figures from previous years are updated in subsequent reports. Years 2012-16 from the 2016 report;[1] years 2008-11 from the 2012 report;[12] years 2002-07 from the 2008 report.[2]
- UK live births are summed from ONS statistics for England and Wales,[15] National Records of Scotland,[16] and Northern Ireland Statistics and Research Agency.[17]
- Incidence estimates are new diagnoses as a proportion of live births; mean age of diagnosis is around 2-3 months though it's not certain all new cases were born in that given year.

Table 24. Genotyped cases per year in the UK CF registry (data from annual reports[1, 2, 12])

	2008	2009	2010	2011	2012	2013	2014	2015	2016
% genotyped	93.7	94.3	95.2	95.6	96.2	97.2	97.7	98.1	98.4
Proportion with CF carrying F508del variant (%)									
Overall	92.0	91.5	91.3	90.6	90.7	90.8	90.4	90.5	90.9
Homozygous F508del	54.3	53.6	52.6	52.0	51.7	51.3	50.6	50.3	50.2
Heterozygous F508del	37.7	38.0	38.7	38.6	39.0	39.5	39.8	40.2	40.7
Other common genotypes (%)									
G551D	5.6	5.8	5.7	5.6	5.6	5.8	5.7	5.6	5.9
R117H	3.2	3.4	3.8	4.1	4.3	4.5	4.5	4.6	5.1
G542X	3.3	3.3	3.5	3.6	3.6	3.6	3.5	3.5	3.6
621+1G→T	2.7	2.7	2.4	2.3	2.1	2.1	2.3	2.2	2.6
N1303K	1.2	1.4	1.4	1.4	1.3	1.3	1.4	1.4	1.7
1717-1G→A	1.2	1.2	1.4	1.3	1.3	1.4	1.3	1.2	1.4
1898+1G→A	1.4	1.2	1.2	1.3	1.1	1.1	1.2	1.2	1.3
"Other"	6.8	8.0	9.2	10.3	12.0	-	16.6	17.5	14.1
Not identified	14.5	13.7	12.9	12.4	10.7	-	6.0	5.3	-

NB No genotype data available pre-2008; "Other" indicates genotype not given in long list. Data is taken from each individual annual report per year.

Table 25. Prevalence of common genotype by nation (2016 CF registry annual report[1])

Variant	England		Scotland		Wales		Northern Ireland	
	N=7890	%	N=823	%	N=433	%	N=398	%
F508del	7149	90.6	745	90.5	392	90.5	335	84.2
G551D	422	5.3	87	10.6	21	4.8	31	7.8
R117H	350	4.4	61	7.4	14	3.2	58	14.6
G542X	233	3.0	59	7.2	24	5.5	24	6.0
621+1G→T	177	2.2	12	1.5	42	9.7	13	3.3
1898+1G→A	89	1.1	<5	-	29	6.7	<5	-

Table 26. Prevalence of mild or severe phenotype (by pancreatic sufficiency)

Study	Design	Overall pancreatic sufficient	By year				Note
			2007	2008	2009	2010	
Hoo et al 2014[13]	Review of UK CF registry data 2007-10. Identification of pancreatic sufficiency based on prescription of pancreatic enzyme replacement therapy. NB: information likely to be contained in the CF registry but is not given in the annual reports	Overall registered cases 2007-10: n=10,516 Pancreatic sufficient 'mild phenotype': 11.7% (n=1235) Pancreatic insufficiency 'severe phenotype': 77.7% (n=8169) Data missing: 10.6% (n=1112)	Cases n=8756 Pancreatic sufficient: 12.8% (n=1235) (Pancreatic insufficient: n=7319) (Data missing: n=358)	Cases n=9004 Pancreatic sufficient: 13.5% (n=1159) (Pancreatic insufficient: n=7417) (Data missing: n=428)	Cases n=9220 Pancreatic sufficient: 13.9% (n=1192) (Pancreatic insufficient: n=7356) (Data missing: n=672)	Cases n=9385 Pancreatic sufficient: 14.4% (n=1220) (Pancreatic insufficient: n=7241) (Data missing: n=924)	There is discrepancy in prevalence figures in annual report (aside from 2010) for unclear reasons. Data on incidence is not given as 43% of new cases did not have data on pancreatic sufficiency so may be unreliable

Criterion 1: question 2: genotype-phenotype association (systematic review)

Table 27. Primary studies on genotype-phenotype association

Study reference	Design and setting	Population characteristics	Genetic comparison	Outcomes
McKone et al 2006[25]	Retrospective cohort US CF Foundation Registry Patients enrolled (observation period) 1993 to 2002 Patients assessed for mortality from time they enrolled (any age) until end of follow-up period	15,651 genotyped and included in the survival model. Participants with one of 21 variants of known functional class and variant frequency >0.1%. Average age at entry 10.2 years for severe and 19.0 years for mild genotypes. Average follow-up 8.6 for severe vs 5.1 years for mild Total registry cohort 30,396 excludes: 2166 with follow-up <1 yr 6877 not genotyped 5702 with unclassified genotype	Severe class vs mild class (both variants class 1-3 vs ≥1 in class 4-5) N=14,525 (93%) severe vs N=1126 (7%) mild class Class 1 (n=2131): G542X, R553X, W1282X, R1162X, 621+1G>T, 1717-1G>A, 1078delT, 3659delC Class 2 (n=11,231): F508del, I507del, N1303K, S549N, G85E Class 3 (n=783): G551D, R560T Class 4 (n=391): R117H, R334W, R347P Class 5 (n=421): 3849+10KbC>T, 2789+5G>A, A455E	1672 deaths during 10 year follow-up (median follow-up 8.6 severe vs 5.1 years mild) Median survival severe: 36.3 years (95% CI 35.5 to 37.6) Median survival mild: 50 years (95% CI 47.1 to 55.9) Among patients who died median age: 24.2 years severe vs 37.6 years mild, p<0.001). <u>Unadjusted analysis association between genotype and survival</u> CFTR genotype (n=15,651): Hazard ratio [HR] 2.25 (95% CI 1.77 to 2.84), p=0.001 *(reported to include patients who died after transplant, but apparently transplant-need not considered as mortality)* <u>Adjusted analyses</u> *Full adjustment for year of entry to the cohort, population size of the CF centre, age, and phenotypic variables of FEV1, BMI, pancreatic sufficiency, and P. aeruginosa colonisation (all documented during year of cohort entry):* CFTR genotype (n=7,305): Hazard ratio [HR] 1.60 (95% CI 1.20 to 2.10), p=0.001 Genotype is an individual predictor of mortality. By phenotype, both FEV1 (HR 0.96, 95% CI 0.96 to 0.97) and BMI (HR 0.88, 95% CI 0.86 to 0.90) were associated with poorer survival, pancreatic insufficiency (PI) had no link and *P. aeruginosa* colonisation a borderline link. Accuracy of CFTR genotype as predictor of age of death using

<table>
<tr><th>Study reference</th><th>Design and setting</th><th>Population characteristics</th><th>Genetic comparison</th><th>Outcomes</th></tr>
<tr><td></td><td></td><td></td><td></td><td>different cut offs:
<table>
<tr><td>Age at death</td><td colspan="4">% (95% confidence interval)</td></tr>
<tr><td></td><td>Sensitivity</td><td>Specificity</td><td>PPV</td><td>NPV</td></tr>
<tr><td><25 years</td><td>98 (97–99)</td><td>8 (6–10)</td><td>53 (51–56)</td><td>81 (70–88</td></tr>
<tr><td><30 years</td><td>98 (97–99)</td><td>11 (8–14)</td><td>69 (67–72)</td><td>71 (60–80</td></tr>
<tr><td><35 years</td><td>97 (96–98)</td><td>14 (10–18)</td><td>82 (80–84)</td><td>57 (46–68</td></tr>
<tr><td><40 years</td><td>97 (96–98)</td><td>20 (14–26)</td><td>91 (89–92)</td><td>44 (33–55</td></tr>
</table>
30 years considered to have the best combination of PPV and NPV as a predictive test:

Of patients who died and had severe genotype, 69% (95% CI 67 to 72%) died before the age of 30.

Of patients who died and had mild genotype, 71% (95% CI 59 to 80%) died after the age of 30.

<u>Reviewer notes</u>
• Overall reliable study
• Test positive (severe genotype): 98% of those who will die before the age of 30 will have severe genotype, but the extremely low specificity demonstrates that many will live beyond this age
• Test negative (mild genotype): 29% of those with mild genotype will die before the age of 30
• Overall shows that genotype would be unreliable for guiding pregnancy decisions based on survival outlook
• Adjusted analysis includes fewer people but genotype is shown as an independent predictor of survival which isn’t</td></tr>
</table>

Study reference	Design and setting	Population characteristics	Genetic comparison	Outcomes
				explained by phenotype alone • Potential that phenotypic markers (e.g. pancreatic enzyme supplements) do not fully capture disease severity though this is used as the assessment method for both studies • Effects of rarer variants other than these 21 are unclear • Uncertain modifying effects of treatment • Prognosis is likely to have improved so this data may not give reliable prognostic information to infants with CF born today
McKone et al 2003[26]	Retrospective cohort US CF Foundation Registry Patients enrolled Jan 1991 to Dec 1999. All patients genotyped at any time during this period assessed to end of follow-up or transplantation (considered as mortality).	N=17,853 genotyped for 23 variants 52% male, 96% white, mean age 10.9 years 1547 deaths (9% of study cohort) Sample genotyped represents 63% of complete registry (n=28,455)	F508del homozygotes vs 11 most common F508del heterozygotes for mortality F508del homozygotes (class 2) vs class of 2nd variant for heterozygotes F508del homozygotes vs 22 F508del heterozygotes for other phenotypic outcomes. All as ACMG 2004 panel with the exception of: S549N included here and not in the 04 panel; 3120+1G>A not included here.	Primary outcome mortality *Crude mortality rate (CMR) per 1000 person-years* *Standardised mortality rate (SMR) by age and sex (95% CI) vs F508del/ F508del* F508del/F508del (n=9144) CMR 19.9; SMR 21.8 (20.5 to 23.1) F508del/G551D (n=593) CMR 18.5; SMR 16.6 (12.4 to 20.8), p=0.019** F508del/G542X (n=574) CMR 17.6; SMR 18.9 (14.1 to 23.7), p=0.257 F508del/N1303K (n=303) CMR 16.9; SMR 16.2 (10.3 to 22.0), p=0.063 F508del/W1282X (n=278) CMR 22.3; SMR 21.6 (14.5 to 28.6), p=0.950 F508del/R553X (n=230) CMR 15.7; SMR 25.0 (11.8 to 38.1), p=0.641 F508del/621+1G>T (n=213) CMR 21.0; SMR 19.2 (11.6 to 26.7) p=0.503

Study reference	Design and setting	Population characteristics	Genetic comparison	Outcomes
				F508del/1717+1G>A (n=120) CMR 21.0; SMR 20.6 (9.9 to 31.4) p=0.833 **F508del/I507del (n=318) CMR 8.9; SMR 8.0 (2.7 to 13.3), p<0.0001** **F508del/R117H (n=177) CMR 9.5; SMR 4.7 (0.8 to 8.5), p<0.0001** **F508del/3849+10kbC>T (n=151) CMR 18.6; SMR 11.9 (5.0 to 18.9), p=0.006** **F508del/2789+5G>A (n=86) CMR 9.0; SMR 4.4 (0.0 to 8.9), p<0.0001** **F508del/other (n=3434) CMR 19.4; SMR 17.6 (15.8 to 19.4), p=0.0002** Other/other (n=2232) CMR 22.2; SMR 20.5 (17.9 to 23.1), p=0.380 * p<0.01 was considered significant for mortality data *Patients needing transplant included as mortality* *By class of second variant vs 2 variants in class 2 (mostly F508del/F508del)* Class 2 (n=9820) CMR 19.6 SMR 21.2 (20.0 to 22.5) Class 1 (n=1670) CMR 19.1 SMR 20.4 (17.4 to 23.4), p=0.615 Class 3 (n=667) CMR 17.6 SMR 16.0 (12.1 to 20.0), p=0.013* Class 4 (n=349) CMR 15.2 SMR 7.8 (4.2 to 11.4), **p<0.0001** Class 5 (n=296) CMR 15.7 SMR 9.1 (4.8 to 13.5), **p<0.0001** Unclassified (n=5051) CMR 20.6 SMR 19.1 (17.4 to 20.7),

Study reference	Design and setting	Population characteristics	Genetic comparison	Outcomes
				p=0.039 *as above p<0.01 considered significant. <u>Secondary outcomes</u> *By class of second variant vs 2 variants in class 2 (mostly F508del/F508del)* Class 2 (n=6599) age at diagnosis 2.6 (+/- 0.1), FEV1 78% predicted (+/- 0.3), FVC 89% predicted (+/- 0.3), PI 92% (91-93), *P. aeruginosa* colonisation 59% (58-61), height 141cm (+/- 0.2), weight 37.0kg (+/- 0.1) Class 1 (n=1158) **age at diagnosis 2.0 (+/- 0.1)**, FEV1 78 (+/- 0.7), FVC 89 (+/- 0.6), PI 91 (90-93), *P. aeruginosa* 59 (56-61), height 140 (+/- 0.4), weight 37.1 (+/- 0.3) Class 3 (n=467) age at diagnosis 3.6 (+/- 0.3), FEV1 77 (+/- 1.1), FVC 89 (+/- 1.1), PI 92 (89-94), *P. aeruginosa* 59 (54-63), height 142 (+/- 0.6), weight 38.3 (+/- 0.5) Class 4 (n=245) **age at diagnosis 11.4 (+/- 0.8), FEV1 85 (+/- 1.4), FVC 94 (+/- 1.2), PI 71 (64-76), *P. aeruginosa* 37 (31-43),** height 143 (+/- 1.2), **weight 41.0 (+/- 1.1)** Class 5 (n=222) **age at diagnosis 12.6 (+/- 0.7),** FEV1 82 (+/- 1.6), FVC 92 (+/- 1.4), **PI 68 (61-74),** *P. aeruginosa* 51 (44-58), height 143 (+/- 1.2), **weight 41.5 (+/- 1.0)** Unclassified (n=3728) **age at diagnosis 6.4 (+/- 0.1), FEV1 81 (+/- 0.4), FVC 90 (+/- 0.4), PI 84 (83-85), *P. aeruginosa* 46 (44-48),** height 141 (+/- 0.2), **weight 38.2 (+/- 0.2)** Outcomes in bold with significance <0.001 vs class 2 (significance level for analysis of phenotypic variables)

Study reference	Design and setting	Population characteristics	Genetic comparison	Outcomes
				Phenotypic variables were mostly collected for year of entry to cohort. *Pancreatic insufficiency assessed by enzyme replacement therapy, P. aeruginosa as positive sputum in past year* *Nutritional status reported for mean 15 year old, 52% male* *By 23 genotypes* F508del/ F508del (n=6213) age at diagnosis 2.5 (+/- 0.1), FEV1 77% predicted (+/- 0.3), FVC 89% predicted (+/- 0.3), PI 92% (91-92), *P. aeruginosa* colonisation 60% (59-61), height 141cm (+/- 0.2), weight 37.0kg (+/- 0.1) *Variables with significant difference p<0.001 from F508del/F508del only:* F508del/G551D (n=411) age at diagnosis 3.7 (+/- 0.3) F508del/I507del (n=149) age at diagnosis 8.5 (+/- 1.1), FEV1 86 (+/- 2.1), PI 84 (78-89), *P. aeruginosa* colonisation 39 (31-48) F508del/R117H (n=123) age at diagnosis 13.7 (+/- 1.2), FEV1 91 (+/- 2.1), FVC 97 (+/- 1.7), PI 65 (55-73), *P. aeruginosa* colonisation 22 (16-29), weight 42.9 (+/- 1.7) F508del/3849+10kB (n=114) age at diagnosis 11.3 (+/- 0.9), PI 66 (57-74), weight 41.2 (+/- 1.2) F508del/2789+5G (n=63) age at diagnosis 13.4 (+/- 1.6), FEV1 88 (+/- 2.8), FVC 97 (+/- 2.3), PI 71 (59-81), *P. aeruginosa* colonisation 32 (22-44) F508del/560T (n=46) FEV1 84 (+/- 3.3) F508del/R347P (n=44) PI 67 (52-79)

Study reference	Design and setting	Population characteristics	Genetic comparison	Outcomes
				F508del/G85E (n=43) age at diagnosis 9.2 (+/- 1.8) F508del/A455E (n=29) age at diagnosis 14.3 (+/- 2.0), FEV1 98 (+/- 4.0), FVC 104 (+/- 3.4), PI 60 (41-76), *P. aeruginosa* colonisation 17 (8-32) F508del/R334W (n=28) age at diagnosis 13.2 (+/- 3.0), PI 67 (46-82) F508del/other (n=2262) age at diagnosis 5.8 (+/- 0.2), FEV1 80 (+/- 0.5), FVC 91 (+/- 0.5), PI 86 (84-87), *P. aeruginosa* colonisation 50 (48-52), weight 38.1 (+/- 0.3) other/other (n=1551) age at diagnosis 7.5 (+/- 0.3), FEV1 82 (+/- 0.6), PI 81 (80-84), *P. aeruginosa* colonisation 40 (38-43), weight 38.3 (+/- 0.3) <u>Reviewer notes</u> Generally shows milder class variants have better outcomes but: • Clear variability within genotype and within same class – couldn't predict outcomes with reliability • Small sample sizes for less common genotypes may limit reliability of analysis • Many participants have uncertain variants and/or those that can't be classified • G85E has since been reclassified from mild to severe class 2 though only 43 were in this group so should have minimal effect • Uncertain confounding including effects of treatment on outcomes and screening on diagnostic age • Potential selection bias, those living longer more likely to

Study reference	Design and setting	Population characteristics	Genetic comparison	Outcomes
				be genotyped • As with most studies – assessment between phenotypic variables and genotype is essentially cross sectional • Also uncertain whether assessment has been the same for all patients across the registry and complete data captured
Lai et al 2004[27]	Retrospective cohort US CF Foundation Registry Patients enrolled 1986 to 2000 and followed up at least once Aim to look at the hypothesis that milder disease presentation have improved prognosis. From this aiming to generate a "baseline risk" component indicating the degree of severity for	N=13,690 for survival, N=3,320 for lung function and N=5290 for *P. aeruginosa* colonisation Total in study of baseline risk with outcome was N=27,703 In total n=18,353 (66.2%) had been genotyped and n=13,690 (49.4%) had one of the classified variants Total potential cohort for period 32,229 with the above excluding: N=2192 with only one-follow-up N=2334 with missing information	F508del/F508del vs 2 severe class variants (1-3) vs ≥1 mild class (4-5) 24 variants listed as: Class 1: G542X, R553X, W1282X, R1162X, 621+1G>T, 1717-1G>A, 1078delT, 3659delC, 2184delA, 2789+5G>A,1898+1G >A, 711+1G>A Class 2: F508del, I507del, N1303K, S549N Class 3: G551D, R560T, A455E Class 4: R117H, R334W, R347P, G85E Class 5: 3849+10KbC>T	Age at diagnosis and disease profile at diagnosis considered as markers of disease severity. Initial analysis examined whether survival/lung function differed according to presentation/diagnostic groups of: • Meconium ileus • Prenatal/newborn screening • Positive family history without symptoms • Symptoms other than meconium ileus Subsequent analysis looked at associations of gender and genotype along with presentation Analysis of the association with genotype comparison to F508del/F508del <u>Shortened survival</u> Severe genotype: odds ratio [OR] 0.76, 95% CI 0.67 to 0.86, **p≤0.001** Mild genotype: OR 0.51, 95% CI 0.37 to 0.70, **p≤0.001** *No definition of "shortened"*

Study reference	Design and setting	Population characteristics	Genetic comparison	Outcomes
	incorporating into survival models	on method of diagnosis		<u>FEV1 <70%</u> Severe genotype: OR 0.88, 95% CI 0.74 to 1.05, ns Mild genotype: OR 1.16, 95% CI 0.55 to 1.33, ns *Assessed by time to first assessment with FEV1 <70%. Analysis of only those aged >6 and with FEV1 >70* <u>*P. aeruginosa* colonisation</u> Severe genotype: OR 1.03, 95% CI 0.95 to 1.11, ns Mild genotype: OR 0.65, 95% CI 0.42 to 1.00 (reported as p≤0.05 though CI is not significant) *Time to first positive culture – but excluding people positive at their first documented visit* <u>Reviewer notes</u> • Exclusion included those with missing data on presentation and follow-up, with small proportion of full cohort with genotyping data • Survival time is unclear other than shortened • Only includes those without *P. aeruginosa* colonisation and with FEV1 >70% at first visit • Only gives the total number for each analysis but doesn't inform how many were in each class (F508del homozygotes vs others with non-F508del severe variants vs mild) • No adjustment for confounding and transplant status not

Study reference	Design and setting	Population characteristics	Genetic comparison	Outcomes
				mentioned • Text lists looking at 25 variants but duplicates printing of R553X • Discrepancy in grouping of G85E, 2789+5G>A, A455E from McKone et al[25, 26] between severe and mild classes; 2184delA, 1898+1G>A and 711+1G>A put has group 1 when other studies have put these are unclassified
O'Connor et al 2002[28]	Retrospective cohort US CF Foundation registry, 1982-1998 (excluding those diagnosed age >18 years) Aim to identify a set of patient/disease characteristics that would be useful for case-mix adjustment for confounders when looking at CF mortality rates	N=15,214 patients and n=1132 deaths Total N=30,469 patients seen during period and N=5906 deaths with others excluded due to lack of socioeconomic and genotyping data.	F508del/F508del vs F508del/other vs other/other (n=8061 homozygotes, n=5414 F508del heterozygotes, n=1829 other)	<u>Multivariate analysis predicting death vs F508del/other</u> F508del/F508del: HR 1.36, 95% CI 1.19 to 1.55, **p<0.001** other/other: HR 1.40, 95% CI 1.15 to 1.71, **p=0.001** Adjustment for: • Gender • Age at diagnosis • Ethnicity (White/Hispanic/Black) • Method of diagnosis/presentation (asymptomatic, respiratory only, GI only, both, meconium/obstruction) • Household income (based on 1990 US Census) <u>Reviewer notes</u> • Uncertain accuracy of data on adjusted variables in registry • Still only 50% genotyped • Unknown identify of other variant: not possible to predict outcome from individual genotypes • Unclear age of death: similarly limited ability to inform

Study reference	Design and setting	Population characteristics	Genetic comparison	Outcomes
				• Potential survivor bias: survivors more likely to be genotyped • Socioeconomic data based on postcode so may be inaccurate
Simmonds et al 2009[29]	Cross sectional survey Case control comparison UK single centre (Royal Brompton Hospital) collecting data since 1965, patients surveyed who'd reached 40 years of age in 2004 Aim to look at hypothesis that older patients within the centre are more likely to have rare and mild genotypes	N=112 aged 40 years vs adult (aged >16 years) CF registry population of 2003 (n=3989) Genotype data available for n=93 (83%) older patients *Patients with transplant included only up to age of transplantation.* Genotyping not described for full adult cohort. State 54% were aged 16-25 years and 17% 35+ but no description of the remaining proportion. Study group: 57% male (vs 56% adult registry). 28% diagnosed in 1st	F508del/F508del vs F508del/other vs other/other n=34 homozygotes n=36 F508del/unknown n=16 F508del/known* n=5 known/unknown n=1 known/known n=1 unknown/unknown * variants reported R117H (n=3), R347P (n=1), G551D (n=4), G542X (n=3), N1303K (n=2), G85E (n=1), 1717-1G>A (n=1) and 621+1G>T (n=1)	<u>Frequency of genotype patients >40 vs remaining registry</u> F508del/F508del: 30% older vs 50% cohort, **p<0.001** F508del/unknown: 32% vs 13%, **p<0.001** F508del/known: 14% vs 22%, p=0.062 Known/unknown: 4% vs 2%, p=0.095 Median age of death 43.1 years for study group, similar to the whole adult registry at 42.8 years. <u>Reviewer notes</u> • Majority of participants in older study group with available data but difficult interpretation due to mix of heterozygotes • Unclear genotyping availability for whole adult population in registry • Indicates that people who live longer are less likely to be F508del homozygotes – but homozygotes still make up 30% of those surviving to 40 so not useful as predictor • Possible interpretation that those living to 40 are more likely to be F508del heterozygotes with rarer variants – though uncertain • Fisher's exact test used because of small between-group

Study reference	Design and setting	Population characteristics	Genetic comparison	Outcomes
		year of life and 32% aged >16, significantly more than whole adult population where 12% were diagnosed >16 (p<0.001) 82% pancreatic insufficiency. No comparison figure for whole adult registry		comparisons • Uncertain confounding from treatment or other factors • Possible survivor bias • Older patients came from a single centre, potentially could have compared within the full registry? • Older patients were also included in the comparison of the full registry though would contribute small number • General finding may still apply but even though this is a UK study, CF care and prognosis is likely very different for those born today compared to 1965
Badet et al 2004[30]	Cross sectional survey Case control analysis French Registry: data collection 1999 of patients born before Jan 1970, diagnosed at <5 years and living to >30 in Dec 1999 Aim: descriptive analysis of	N=114 aged 30 years vs total CF registry population of 1999 (n=3220) Genotype data available for 105 (92%) older patients, both variants identified in 93 (82%) vs 79% of the full registry Mean age of "survivors" 34.3 years with mean age at diagnosis 28.3 months.	F508del/F508del vs F508del/other vs other/other Of n=31 F508del/other the most frequent were reported as 2789+5G>A (n=4), G542X (n=4), R347H (n=3), 1717-1G>A (n=2) and R553X (n=2)	<u>Frequency of genotype in patients >30 years vs remaining registry</u> F508del/F508del: 56% older vs 58% cohort F508del/other: 33% vs 21% other/other: 11% vs 21% Reported no significant differences (P>0.05) <u>Reviewer notes</u> • Majority of participants genotyped in both groups but difficult interpretation and unclear what all second variants were • Conflicting with other studies it finds no differences in proportions between F508del/F508del and F508del/other

Study reference	Design and setting	Population characteristics	Genetic comparison	Outcomes
	those with long-term survival and follow-up			• Very limited information on statistical analysis and small groupings which may be unreliable • Unclear effect of confounding from treatment or other factors • May not be representative of all older patients as excluding those diagnosed >5 years • As Simmonds et al, older patients may also have been included in the comparison of the full registry though would contribute small number • CF care and prognosis is likely very different for those born today compared to <1970
Koch et al 2001[31]	Cross sectional European Epidemiologic Registry of CF (ERCF) including data on patients from 9 European countries (Belgium, Denmark, France, Germany, Ireland, The Netherlands, Sweden, Austria, and UK) since	N=8963 with observation period of 180 days from time of enrolment (though first phenotypic assessment from time of enrolment taken) Representing 76% genotyped of the total 11,749 in registry –thought to represent half the patients across the 9 countries Patients stratified into >18 and <18 years	By class Patients grouped according to functional class combination of the 2 variants Initially grouped as homozygotes but as there were few homozygotes for 3, 4 or 5 variants, some groups were pooled to give final for analysis : • class 1/1 • class 2/2 • class 2/3 • class 3/3	<u>Assessed population mean age</u> *First input in registry* • 1/1 (n=72) 10.7 years (95% CI 9.0 to 12.4) • 2/2 (n=5020) 12.4 years (95% CI 12.1 to 12.6) • 2/3 (n=265) 13.4 years (95% CI 12.4 to 14.4) • 3/3 (n=23) 15.6 years (95% CI 11.7 to 19.5) • 4/any (n=187) 16.0 years (95% CI 14.4 to 17.6) • 5/any (n=22) 17.0 years (95% CI 12.7 to 21.4) <u>Weight for age percentile</u> *First valid value in registry* <18 years • 1/1 (n=60) 25.9 (95% CI 19.5 to 32.4)

Study reference	Design and setting	Population characteristics	Genetic comparison	Outcomes
	1994. Data accessed in August 1997. Aim to look at disease manifestations by class		• class 4/any • class 5/any 87 variants analysed (of 187):	• 2/2 (n=3738) 32.4 (95% CI 31.5 to 33.3) • 2/3 (n=194) 39.0 (95% CI 34.9 to 43.1) • 3/3 (n=12) 29.7 (95% CI 12.6 to 46.8) • 4/any (n=121) 42.3 (95% CI 37.0 to 47.6) • 5/any (n=13) 36.6 (95% CI 20.4 to 52.7) ≥18 years • 1/1 (n=12) 14.0 (95% CI 2.8 to 25.3) • 2/2 (n=1273) 26.8 (95% CI 25.4 to 28.2) • 2/3 (n=70) 25.0 (95% CI 19.4 to 30.6) • 3/3 (n=11) 22.2 (95% CI 13.1 to 31.4) • 4/any (n=63) 44.3 (95% CI 36.9 to 51.6) • 5/any (n=9) 15.6 (95% CI 3.6 to 27.5) <u>Pancreatic insufficiency</u> *Enzyme replacement therapy but unclear time of assessment* <18 years • 1/1 (n=58/60) 96.7% (95% CI 88.5 to 99.6) • 2/2 (n=3670/3744) 98.0% (95% CI 97.5 to 98.4) • 2/3 (n=187/194) 96.4% (95% CI 92.7 to 98.5) • 3/3 (n=11/12) 91.7% (95% CI 61.5 to 99.8) • 4/any (n=87/122) 71.3% (95% CI 62.4 to 79.1) • 5/any (n=8/13) 61.5% (95% CI 31.6 to 86.1)

Study reference	Design and setting	Population characteristics	Genetic comparison	Outcomes
				≥18 years • 1/1 (n=12/12) 100% (95% CI 73.5 to 100) • 2/2 (n=1245/1276) 97.6% (95% CI 96.6 to 98.3) • 2/3 (n=68/71) 95.8% (95% CI 88.1 to 99.1) • 3/3 (n=11/11) 100% (95% CI 71.5 to 100) • 4/any (n=34/65) 52.3% (95% CI 39.5 to 64.9) • 5/any (n=4/9) 44.4% (95% CI 13.7 to 78.8) <u>FEV1 % predicted</u> *First valid value in registry* <18 years • 1/1 (n=33) 71.3 (95% CI 64.3 to 78.3) • 2/2 (n=1973) 76.4 (95% CI 75.3 to 77.6) • 2/3 (n=121) 78.9 (95% CI 74.7 to 83.1) • 3/3 (n=5) 65.1 (95% CI 33.3 to 96.9) • 4/any (n=73) 82.8 (95% CI 78.2 to 87.4) • 5/any (n=13) 75.2 (95% CI 64.7 to 85.8) ≥18 years • 1/1 (n=11) 50.2 (95% CI 33.7 to 66.7) • 2/2 (n=1032) 54.2 (95% CI 52.7 to 55.6) • 2/3 (n=57) 58.0 (95% CI 51.1 to 64.9) • 3/3 (n=10) 60.8 (95% CI 42.3 to 79.2)

Study reference	Design and setting	Population characteristics	Genetic comparison	Outcomes
				• 4/any (n=48) 61.8 (95% CI 54.1 to 69.5) • 5/any (n=8) 45.9 (95% CI 31.3 to 60.6) <u>FVC %</u> *First valid value in registry* <18 years • 1/1 (n=33) 86.9 (95% CI 80.4 to 93.3) • 2/2 (n=1970) 85.5 (95% CI 84.3 to 86.8) • 2/3 (n=121) 88.3 (95% CI 83.7 to 92.9) • 3/3 (n=5) 78.9 (95% CI 52.6 to 105.3) • 4/any (n=73) 89.4 (95% CI 84.9 to 94.0) • 5/any (n=13) 83.3 (95% CI 75.2 to 91.3) ≥18 years • 1/1 (n=11) 67.4 (95% CI 54.3 to 80.5) • 2/2 (n=1032) 71.8 (95% CI 70.3 to 73.2) • 2/3 (n=57) 74.1 (95% CI 67.4 to 80.7) • 3/3 (n=10) 73.2 (95% CI 55.7 to 90.6) • 4/any (n=48) 76.5 (95% CI 69.8 to 83.2) • 5/any (n=8) 71.4 (95% CI 55.0 to 87.7) <u>*P. aeruginosa* colonisation</u> *Unclear assessment*

Study reference	Design and setting	Population characteristics	Genetic comparison	Outcomes
				<18 years • 1/1 (n=29/58) 50.0% (95% CI 36.6 to 63.4) • 2/2 (n=1767/3537) 50.0% (95% CI 48.3 to 51.6) • 2/3 (n=103/187) 55.1% (95% CI 47.7 to 62.3) • 3/3 (n=5/9) 55.6% (95% CI 21.2 to 86.3) • 4/any (n=38/115) 33% (95% CI 24.6 to 42.4) • 5/any (n=11/12) 91.7% (95% CI 61.5 to 99.8) ≥18 years • 1/1 (n=12/12) 100% (95% CI 73.5 to 100) • 2/2 (n=1019/1239) 82.2% (95% CI 80.0 to 84.3) • 2/3 (n=58/71) 81.7% (95% CI 70.7 to 89.9) • 3/3 (n=10/10) 100% (95% CI 69.2 to 100) • 4/any (n=34/60) 56.7% (95% CI 43.2 to 69.4) • 5/any (n=9/9) 100% (95% CI 66.4 to 100) <u>Reviewer notes</u> • Groups 3/3 and 5/any too small for reliable comparison • Clear pattern that people with class 1/1, 2/2 or 2/3 variants were younger • Those with class 4/any variants clearly had higher weight for age percentiles and were less likely to have pancreatic insufficiency than those with class 1/1, 2/2 or 2/3, generally without overlapping confidence intervals • Mean lung function parameters slightly higher in class

Study reference	Design and setting	Population characteristics	Genetic comparison	Outcomes
				4/any which may be associated with the finding of less *P. aeruginosa* colonisation • Potential for misclassification: some added based on similar localisation within the gene. Also others included by McKone[25] and de Boeck[44] listings have subsequently been reclassified (G85E subsequently changed from class 4 to 2) • Unclear assessment for all variables • Statistical analysis reportedly not performed as the study was not hypothesis-testing • No adjustment for confounders and care and treatment may not be applicable to present UK
Dewulf et al 2015[24]	Retrospective cohort Belgian CF Registry Patients enrolled 2010 *No screening*	N=748 Total registry n=1138 of whom n=853 (75%) had known variants that could be classified Additional 105 patients excluded because they'd received a transplant gives final sample of n=748 (66% of original cohort)	Severe class vs mild class (both variants class 1-3 vs ≥1 in class 4-5) N=759 severe vs N=94 mild Analysed variants not listed but classes said to comply with de Boeck et al[44]	<u>Treatments used for >3 months in that year</u> Grouped according to 3 categories: • Low: inhaled bronchodilators, inhaled corticosteroids, oral antibiotics • Medium: inhaled antibiotics, enzyme therapy, hypertonic saline, mucolytics, oral corticosteroids • High: IV antibiotics, parenteral nutrition, oxygen, gastrostomy, insulin Weighted treatment burden index (TBI) assessed by multiplying number of therapies in each by, respectively, 1, 2 or 3: • TBI: severe 9 (IQR 6-12) vs mild 6 (IQR 3-8), p<0.001 Regression analysis of TBI: • Significant effect of mild vs severe class (Exp 0.7685,

Study reference	Design and setting	Population characteristics	Genetic comparison	Outcomes
				95% CI 0.6952 to 0.8495, p <0.001) • Mild class associated with 23.1% decrease in treatment (95% CI 15.0 to 30.5) *Adjustment for age, gender, FEV1* By number of therapies: • Median 5 (IQR 4-7) for severe vs 4 (2-5) for mild class • Low intensity: 2 (IQR 2-3) vs 2 (IQR 1-3) • Medium intensity: 2 (IQR 2-3) vs 2 (IQR 1-2) • High intensity: 1 (IQR 0-1) vs 0 (IQR 0-1) All p<0.001 • Hospitalised patients: 50.8% vs 24.7%, p<0.001 (no difference in hospitalised days for this subgroup) • Receipt of IV antibiotics: 46.0% vs 23.5%, p<0.001 <u>Other outcomes</u> Age at diagnosis: Severe 0.3 years (IQR 0.1 to 1.3) vs mild 5.2 years (IQR 0.4 to 20.9), **p<0.001** Pancreatic insufficient: severe n=655 (98.8%) vs mild n=31 (36.5%), **p<0.001** FEV1: severe 77.0% predicted (IQR 55.6 to 94.1) vs mild 86.8% (IQR 68.1 to 103.0), **p<0.001** Chronic *P. aeruginosa* infection: severe n=240 (36.2%) vs mild n=12 (14.1%), **p<0.001**

Study reference	Design and setting	Population characteristics	Genetic comparison	Outcomes
				Chronic P. aeruginosa infection using defined criteria. Pancreatic sufficiency assessed by fat loss in stool and fecal elastase Reviewer notes • Recent study in non-screening setting • Treatment burden estimate only based on data in registry • Pancreatic sufficiency assessed by fat loss in stool and fecal elastase which may give better precision though this is less comparative against other studies looking at pancreatic enzyme replacement • Better genotyping availability but still only 75% genotyped • No adjustment for confounders for all phenotypic variables
Green et al 2010[32]	Retrospective cohort US CF Twin and Sibling Study (CFTSS) Followed after enrolment (date not given to Dec 2008) Aim to assess the correlation between	N=1381 all participants having a twin or sibling with CF and having sputum culture Representing 83% of original cohort of n=1659 excluding n=35 with no infection data, n=16 with no genotype data and n=227 whose variants couldn't be classified.	Severe class (both variants 1-3) vs mild (≥1 variant 4-5)	Infection Assessment of 13 bacterial strains, 9 of which were associated with higher prevalence in severe classes. Analysis performed for *P. aeruginosa* (Pa), mucoid Pa (MPa), and Aspergillus fumigatus (Asp) using four criteria: • first positive culture with organism (previous negative culture a minimum of 1 week prior) • chronic infection: 3 positive cultures within 6 months with each culture separated by at least 1 month chronic infection (similar to European criteria) • multiple infection: at least 3 positive cultures, but not meeting the definition for chronic infection (as most patients in the US do not attend CF clinic 3 times in 6

<table>
<tr><th>Study reference</th><th>Design and setting</th><th>Population characteristics</th><th>Genetic comparison</th><th>Outcomes</th></tr>
<tr><td></td><td>CFTR functional genotype and infection with a variety of pathogens using detailed infection data from this cohort</td><td></td><td></td><td>months)
• persistent infection: multiple cultures obtained in 3 consecutive years with positive cultures observed in at least 2 of the 3 years (said to be used in a recent CF modifier study)

Earlier age of acquisition of Pa for severe class (5.5 years) vs mild class (14.5 years; p<0.001)

Risk of Pa higher for severe class than mild class variants by all definitions:

<table>
<tr><td>Definition</td><td>Total positive</td><td>Severe (% positive)</td><td>Mild (% positive)</td><td>aHR (95% CI)</td></tr>
<tr><td>First</td><td>318 (436)</td><td>278 (79.4%)</td><td>40 (46.5%)</td><td>3.17 (2.10 to 4.78)</td></tr>
<tr><td>Chronic</td><td>127 (436)</td><td>118 (33.7%)</td><td>9 (10.5%)</td><td>5.47 (2.20 to 13.58)</td></tr>
<tr><td>Multiple</td><td>229 (436)</td><td>206 (58.9%)</td><td>23 (26.7%)</td><td>3.81 (2.32 to 6.28)</td></tr>
<tr><td>Persistent</td><td>228 (436)</td><td>203 (58.0%)</td><td>25 (29.1%)</td><td>3.32 (2.00 to 5.50)</td></tr>
</table>

Results similar for MPa and Asp.
Adjusted for FEV1 in the year period to first infection and number of cultures performed.
(Gender, ethnicity and pancreatic status had also initially been assessed in univariate regression analysis but as they didn’t have significant effect weren’t included in the final model.)

<u>Baseline characteristics</u></td></tr>
</table>

Study reference	Design and setting	Population characteristics	Genetic comparison	Outcomes
				• Pancreatic insufficiency: severe 97.8% (n=1180) vs mild 30.3% (n=46), p<0.001 • Max FEV1 since last clinic visit: severe 0.68 ±0.26 (n=1111) vs 0.75 ±0.25 (n=145), p<0.001 • Average cultures per year: severe 3.89 ±1.92 (n=1201) 3.93± 2.27 (n=159), ns <u>Reviewer notes</u> • Valuable use of different definitions of infection, suggests positive effect is independent of criteria used • Calculated adequate sample size with 80% power • May not be representative as all participants had to have a surviving sibling with CF for inclusion • Reportedly 85% of study sample had Pa colonisation compared with 53% of the CF Registry in 2008. As all had to have had sputum culture for inclusion, unclear whether they may be representative • Other phenotypic measures are characteristics of the sample but not the primary aim of the study – unclear how pancreatic status was defined • Some potential for misclassification; based on McKone et al 2003[26] though with various additional

<table>
<tr><th>Study reference</th><th>Design and setting</th><th>Population characteristics</th><th>Genetic comparison</th><th>Outcomes</th></tr>
<tr>
<td>Radtke et al 2017[33]</td>
<td>Cross sectional

International, multicentre members of the Exercise Working Group of the European CF Society

17 centres: 3 in Canada, 2 in US, and remainder in UK (n=39 patients), Australia, Austria, France, Germany, Greece, Israel, Netherlands, Italy, Spain, Serbia.

Representative of 32 centres asked to provide data on ≥20 patients, aged ≥8 years who completed a</td>
<td>N=726

Represents 73% of total potential sample of n=990.

n=119 with missing genotype data (88%) available with further exclusions n=120 with missing cardiopulmonary exercise testing (CPET), n=12 aged <8 years and remainder with other missing data.

Sample 45% female, average 18.7 years (range 8 to 61 years), FEV1 76.6 +/- 22.9</td>
<td>Severe class (both variants 1-3) vs mild (≥1 variant 4-5)

Classification by geneticist blinded to exercise data.</td>
<td><u>Characteristics by class of milder of 2 variants</u>
<table>
<tr><td>Highest class</td><td>1/1</td><td>≤2/2</td><td>≤3/3</td><td>≤4/4</td><td>≤5/5</td></tr>
<tr><td>Number patients</td><td>32</td><td>550</td><td>39</td><td>63</td><td>42</td></tr>
<tr><td>VO_2 peak L/min</td><td>1.6 (1.3-1.8)</td><td>1.7 (1.4-2.3)</td><td>1.8 (1.3-2.2)</td><td>1.8 (1.5-2.3)</td><td>1.7 (1.3-2.4)</td></tr>
<tr><td>$Watt_{max}$</td><td>111 (83-140)</td><td>127 (98-170)</td><td>130 (95-163)</td><td>124 (95-170)</td><td>130 (85-180)</td></tr>
<tr><td>FEV1 (% predicted)</td><td>80 (45-93)</td><td>79 (60-94)</td><td>78 (50-90)</td><td>86 (72-96)</td><td>80 (62-9</td></tr>
<tr><td>BMI (kg/m^2) ♥</td><td>18.8 (16.9-20.1)</td><td>19.3 (17.3-21.5)</td><td>20.4 (17.5-24.2)</td><td>20.6 (18.8-23.0)</td><td>22.3 (19 25.0)</td></tr>
<tr><td>Body fat (%) ♥</td><td>17.2 +/- 14.7</td><td>18.2 +/- 5.7</td><td>19.9 +/- 5.5</td><td>21.4 +/- 6.4</td><td>22.4 +/-</td></tr>
<tr><td>Pancreatic insufficient ♥</td><td>97%</td><td>93%</td><td>89%</td><td>24%</td><td>24%</td></tr>
<tr><td>P. aeruginosa ♦</td><td>100%</td><td>95%</td><td>55%</td><td>37%</td><td>36%</td></tr>
</table>
♥ p<0.001 ♦ p<0.05 for difference between groups
<table>
<tr><td>Class</td><td>Both variants 1-3</td><td>≥1 variant 4-5</td></tr>
<tr><td>Number patients</td><td>621</td><td>105</td></tr>
<tr><td>VO_2 peak L/min</td><td>1.74 (1.4-2.2)</td><td>1.78 (1.4-2.4)</td></tr>
<tr><td>$Watt_{max}$</td><td>125 (95-168)</td><td>130 (94-176)</td></tr>
<tr><td>FEV1 (% predicted)</td><td>79 (59-93)</td><td>84 (68-96)</td></tr>
<tr><td>BMI z score ♦ (kg/m^2 not given)</td><td>-0.25 (-0.95 to 0.42)</td><td>-0.11 (-0.77 to 0.74)</td></tr>
</table>
</td>
</tr>
</table>

<table>
<tr><th>Study reference</th><th>Design and setting</th><th>Population characteristics</th><th>Genetic comparison</th><th>Outcomes</th></tr>
<tr><td></td><td>maximal cardiopulmonary exercise test (CPET) between January 1999 and December 2014.

States contacting study centres

Aim to investigate factors associated with peak oxygen uptake (VO2 primary outcome) and maximum work rate (Wattmax), focusing on genotype functional class</td><td></td><td></td><td>
<table>
<tr><td>Body fat (%)♥</td><td>18.2 +/- 5.7</td><td>21.8 +/- 6.4</td></tr>
<tr><td>Pancreatic insufficient ♦</td><td>95%</td><td>24%</td></tr>
<tr><td>P. aeruginosa ♥</td><td>54%</td><td>36%</td></tr>
</table>

Mixed models adjusted for age, sex, BMI z score, FEV1 and P. aeruginosa found no effect of CFTR group on main outcomes:
• VO_2 % predicted: ß coefficient -0.95 (-4.18 to 2.29), p=0.57
• $Watt_{max}$ % predicted: ß coefficient -1.38 (-5.04 to 2.27), p=0.46

P. aeruginosa assessed by at least 2 of 4 samples positive in past year. Assessment of pancreatic sufficiency unclear.

<u>Reviewer notes</u>
• Doesn’t support a role of genotype class on exercise capacity
• Exercise capacity primary outcome rather than other variables
• Uncertain assessment of all other phenotypic variables
• Small groups for individual group analyses
• Recognised class system based on McKone et al[25] but with extra additions
• Participants predominantly with milder lung function
• Adjustment for other factors only in analysis of exercise capacity and no other variables</td></tr>
</table>

Study reference	Design and setting	Population characteristics	Genetic comparison	Outcomes
				• Recent large international sample but there may be differences in treatment and care
The Cystic Fibrosis Genotype-Phenotype Consortium [34] 1993	Cross sectional time period of analysis unclear 32 of 89 centres belonging to the CF Genetic Analysis Consortium 22 centres genotyped 100% of patients, 2 genotyped 75%, 8 didn't specify Aim to describe clinical features of	N=399 F508del/F508del N=399 F508del/other Homozygotes age- and sex-matched against those with the next 7 most common genotypes. Participants matched within centre to limit variation in treatment	F508del/F508del vs F508del/other: G542X, R553X, W1282X, N1303K, R117H, 621+1G>T, 1717-1G>A	<u>Variables by genotype</u> F508del/F508del (n=399) current age 13.0 (+/- 8.7), age at diagnosis 1.7 (+/- 3.0), FEV1 70% predicted (+/- 27), PS (pancreatic sufficiency) 2.5%, *P. aeruginosa* colonisation 56%, Shwachman clinical score 75 F508del/G542X (n=148) current age 11.9 (+/- 8.7), age at diagnosis 1.6 (+/- 3.1), FEV1 67% predicted (+/- 27), PS 0♦, *P. aeruginosa* 42%, Shwachman score 74 F508del/R553X (n=52) current age 12.5 (+/- 8.1), age at diagnosis 1.7 (+/- 2.7), FEV1 64% predicted (+/- 25), PS 2%, *P. aeruginosa* 66%, Shwachman score 79 F508del/N1303K (n=60) current age 12.3 (+/- 8.0), age at diagnosis 1.5 (+/- 2.7), FEV1 69% predicted (+/- 24), PS 0, *P. aeruginosa* 53%, Shwachman score 72 F508del/W1282X (n=17) current age 11.0 (+/- 10.8), age at diagnosis 4.0 (+/- 9.9), FEV1 75% predicted (+/- 26), PS 0, *P. aeruginosa* 82%, Shwachman score 79 F508del/1717-1G>A (n=30) current age 11.8 (+/- 8.0), age at diagnosis 2.0 (+/- 4.4), FEV1 68% predicted (+/- 26), PS 3%, *P. aeruginosa* 48%, Shwachman score 71 F508del/621+1G>T (n=51) current age 14.6 (+/- 7.7)♦, age at diagnosis 0.8 (+/- 1.1), FEV1 73% predicted (+/- 26), PS 2%, *P.*

<table>
<tr><th>Study reference</th><th>Design and setting</th><th>Population characteristics</th><th>Genetic comparison</th><th>Outcomes</th></tr>
<tr><td></td><td>F508del compound heterozygotes with other common genotypes</td><td></td><td></td><td>
aeruginosa 63%, Shwachman score 75

F508del/R117H (n=23) current age 23.5 (+/- 9.6), age at diagnosis 10.2 (+/- 10.5), FEV1 73% predicted (+/- 22), PS 87%, P. aeruginosa 30%, Shwachman score 81

♦ p=0.03 vs F508del homozygotes though expected to be a chance finding

Shwachman clinical score looks at general activity, physical examination, nutrition and radiological findings. Score: excellent (86-100), good (71-85), average (56-70), poor (41-55) or severe (≤40)

Lung function: most centres reported the best of 3 efforts in one day

P. aeruginosa first positive culture after series of negative. Routinely performed at every 3-monthly clinic visit.

Variable assessment of pancreatic sufficiency across centres.

F508del/R117H significantly different from F508del homozygotes.

Specific age- and sex-matched comparison of 23 pairs:
<table>
<tr><td></td><td>F508del/F508del</td><td>F508del/R117H</td><td>p</td></tr>
<tr><td>Age at diagnosis (years)</td><td>2.5 (+/- 4.3)</td><td>10.2 (+/- 10.5)</td><td>0.002</td></tr>
<tr><td>FEV1 (% predicted)</td><td>69 (+/- 23)</td><td>73 (+/- 22)</td><td>0.5</td></tr>
<tr><td>Pancreatic sufficient</td><td>4% (1/23)</td><td>87% (20/23)</td><td><0.001</td></tr>
<tr><td>Shwachman score</td><td>77 (+/- 14)</td><td>84 (+/- 11)</td><td>0.07</td></tr>
</table>

<u>Reviewer notes</u>
</td></tr>
</table>

<table>
<tr><th>Study reference</th><th>Design and setting</th><th>Population characteristics</th><th>Genetic comparison</th><th>Outcomes</th></tr>
<tr><td></td><td></td><td></td><td></td><td>

Indicates R117H as a variant conferring pancreatic sufficiency
Cross sectional sample representing only one third of centres
Time period unknown
Unclear where samples with different genotype were recruited from, e.g. whether they included all people with this genotype across all centres
Small size of genotype groups for comparison decreases reliability of analyses
Variable methods used to define pancreatic status
No adjustment for confounding aside from age, sex and treatment centre

</td></tr>
<tr><td>Szczesniak et al 2017[36]</td><td>Retrospective cohort

US CF Foundation Patient Registry

Patients with FEV1 data recorded when aged 6-21 years between Jan 1997 and Dec 2013

Aim to identify and</td><td>N=18,387 with median 19 FEV1 observations each over 6.8 years of follow-up

Decline in FEV1 was assessed by a functional data analysis technique known as functional principal components analysis for sparse longitudinal data (FPCA)

Patients grouped by phenotype</td><td>Given as number of F508del copies:

2 (corresponding to F508del/F508del) vs 1 (F508del/other) vs none (other/other)</td><td><u>Genotype data on rate of decline in FEV1</u>
<table>
<tr><th></th><th>Total patients</th><th colspan="3">Grouped by decline in FEV1</th></tr>
<tr><th></th><th></th><th>Early</th><th>Middle</th><th>Late</th></tr>
<tr><td>F508del/F508del</td><td>6,013 (32.7%)</td><td>1,347 (29.3%)</td><td>3,062 (33.3%)</td><td>1,586 (34.5%)</td></tr>
<tr><td>F508del/other</td><td>8,568 (46.6%)</td><td>2,055 (44.7%)</td><td>4,321 (47.0%)</td><td>2,188 (47.6%)</td></tr>
<tr><td>other/other</td><td>3,806 (20.7%)</td><td>1,195 (26.0%)</td><td>1,811 (19.7%)</td><td>822 (17.9%)</td></tr>
</table>
Overall trend for number of copies F508del given as p<0.0001

Genotype as predictor of early decline in FEV1: comparison to F508del homozygotes:

F508del/other: Odds Ratio 0.99 (95% CI 0.80 to 1.23),

</td></tr>
</table>

<table>
<tr><th>Study reference</th><th>Design and setting</th><th>Population characteristics</th><th>Genetic comparison</th><th>Outcomes</th></tr>
<tr><td></td><td>characterize phenotypes of rapid FEV1 decline for adolescents and young adults with CF, and to identify phenotypic predictors of earlier rapid FEV1 decline</td><td>according to pattern of FEV1 decline:
Early (< quartile 1 on FPCA): continual loss with max decline 3.2% per year at 12.9 years
Middle (quartile 1-3): max loss 2.8% per year at 16.3 years
Late (> quartile 3): max loss 2.9% per year at 18.5 years</td><td></td><td>p=0.0016*
Other/other: OR 1.73 (95% CI 1.36 to 2.21), p<0.0001
*apparently significant p value but confidence intervals span zero; no discussion of genotyping results
Apparent adjustment for other baseline predictors of gender, age at diagnosis, birth cohort year, socioeconomic status and phenotyoic variables.
<u>Reviewer notes</u>
Large recent sample aiming is to model baseline predictors of decline (including birth year, age at diagnosis, BMI pancreatic status, infections, diabetes, socioeconomics)
All have genotyping data so unclear how representative they may be of the initial registry sample
No discussion of genotyping results and limited information can be drawn, for example, suggesting that people carrying non-specific variants other than F508del will have early rapid decline in lung function
Uncertain significance around p values</td></tr>
<tr><td>De Boeck and Zolin 2017[37]</td><td>Retrospective cohort
European CF Society Patient Registry (ECFSPR) containing information from 15</td><td>N=11,417 patients aged >6 years of age without lung transplant and with lung function data collected in ≥2 years
Total eligible n=35,259</td><td>F508del/F508del vs heterozygotes with variant combinations of:
Class 1* and class 1/2
Class 3 and class 1/2/3</td><td><u>FEV1</u>
<table>
<tr><th rowspan="2">Genotype group</th><th colspan="3">Proportion (%) with FEV1 predicted</th><th rowspan="2">Mean annual change in FEV1 (95%CI)</th></tr>
<tr><th><40% (n=1349)</th><th>40-90% (n=6964)</th><th>>90% (n=3104)</th></tr>
<tr><td>F508del/F508del (n=8152)</td><td>12</td><td>61</td><td>26</td><td>-1.52 (-1.72 to -1.31)</td></tr>
<tr><td>≥ class 1 (n=1959)</td><td>11</td><td>61</td><td>28</td><td>-1.35 (-1.70 to -0.99)</td></tr>
</table></td></tr>
</table>

<table>
<tr><th>Study reference</th><th>Design and setting</th><th>Population characteristics</th><th>Genetic comparison</th><th>Outcomes</th></tr>
<tr>
<td></td>
<td>registries and 50 centres across 12 countries

Data on patients from 2008, 09 and 10

Aim to look at yearly change in FEV1 according to variant class</td>
<td>n=33,820 had DNA analysis

n=32,329 had at least one variant identified

n=21,608 could be classified into the assessment groups

Further exclusions due to age <6 years (n=4304), receipt of transplant or no data on transplant status (n=1224), and <2 FEV1 measures (n=4663).</td>
<td>• Class 4 and class 1/2/4
• Class 5 and class 1/2/5

*a class 1 stop codon variant (which has been treated with the off-label treatment ataluren): list of specific variants not given</td>
<td>
<table>
<tr><td>≥ class 3 (n=553)</td><td>13</td><td>62</td><td>25</td><td>-1.24 (-1.87 to -0.61)</td></tr>
<tr><td>≥ class 4 (n=463)</td><td>7</td><td>54</td><td>39</td><td>-0.62 (-1.30 to +0.06)</td></tr>
<tr><td>≥ class 5 (n=290)</td><td>9</td><td>60</td><td>31</td><td>-0.35 (-1.21 to +1.0)</td></tr>
</table>
Having ≥ one variant class 4 or 5 confers better lung function than other groups (p<0.00001)

No difference between groups found for annual change but pooling class 4 and 5 found small difference of +0.88% in yearly change compared to the other three groups (p=0.004)

Similar results on analysis of those with FEV1 40-90%, specifically.

Analysis of those with FEV1 >90% revealed that change was greatest in these patients, and markedly different for class 4 and 5 compared to the other groups (p not given):
• F508del/F508del: -4.00 (-4.66 to -3.33)
• ≥class 1: -4.28 (-5.15 to -3.40)
• ≥class 3: -4.28 (-5.71 to -2.85)
• ≥class 4: -1.88 (-3.07 to -0.69)
• ≥class 5: -1.78 (-3.44 to -0.12)

Adjustment for age only

<u>Reviewer notes</u>
• Large European registry analysis with recent data but only a third of potential participants genotyped and</td>
</tr>
</table>

<table>
<tr><th>Study reference</th><th>Design and setting</th><th>Population characteristics</th><th>Genetic comparison</th><th>Outcomes</th></tr>
<tr><td></td><td></td><td></td><td></td><td>analysed
• Possible selection bias from age group and those having repeat assessments though similar frequency of classes when excluding on this basis which gives greater confidence in the findings
• Confounders not assessed with likely differences in care and screening practices across Europe
• Variant class reliable but analysed groups exclude certain class combinations</td></tr>
<tr><td>Duguepero ux and De Braekeleer 2005[35]</td><td>Cross sectional
French CF registry patients who attended participating centres 1992 to 2002 and carrying variants 3849+10kbC>T or 2789+5G>A.
Analysis of F508del heterozygotes seen during year 2000.

Exclusion of</td><td>N=16 with genotype 3849+10kbC>T/F508del age and sex-matched to n=16 F508del/F508del
N=34 with genotype 2789+5G>A /F508del age and sex-matched to n=34 F508del/F508del
Matched pairs came from the same centre
Of total n=38 carrying 3849+10kbC>T and n=82 carrying 2789+5G>A – exclusion of heterozygotes other than F508del and</td><td>Specific genotype comparison
3849+10kbC>T/F508del
2789+5G>A/F508del
vs F508del/F508del</td><td><u>Phenotypic comparison</u>
3849+10kbC>T/F508del vs F508del/F508del
<table>
<tr><th>Mean values</th><th>3849+10kbC>T/F508del (n=16)</th><th>F508del/F508del (n=16)</th><th>P value</th></tr>
<tr><td>Age at diagnosis</td><td>12.7 +/- 9.6</td><td>3.1 +/- 5.1</td><td>0.002</td></tr>
<tr><td>Pancreatic insufficient %</td><td>46.6</td><td>100</td><td>0.002</td></tr>
<tr><td>FEV1 % predicted</td><td>83.04 +/- 12.08</td><td>59.86 +/- 21.11</td><td>0.069</td></tr>
<tr><td>FVC % predicted</td><td>91.60 +/- 8.19</td><td>76.96 +/- 20.80</td><td>0.082</td></tr>
<tr><td>BMI kg/m^2</td><td>16.28 +/- 3.26</td><td>16.11 +/- 3.00</td><td>Not significant</td></tr>
</table>
Pancreatic status variable assessment

2789+5G>A/F508del vs F508del/F508del
<table>
<tr><th>Mean values</th><th>2789+5G>A/F508del (n=34)</th><th>F508del/F508del (n=34)</th><th>P value</th></tr>
<tr><td>Age at diagnosis</td><td>16.6 +/- 12.7</td><td>4.5 +/- 8.9</td><td>0.0001</td></tr>
</table></td></tr>
</table>

<table>
<tr><th>Study reference</th><th>Design and setting</th><th>Population characteristics</th><th>Genetic comparison</th><th>Outcomes</th></tr>
<tr><td></td><td>those screened</td><td>those not seen in 2000.</td><td></td><td>
<table>
<tr><td>Pancreatic insufficient %</td><td>59.4</td><td>97.0</td><td>0.002</td></tr>
<tr><td>FEV1 % predicted</td><td>75.38 +/- 29.69</td><td>59.06 +/- 24.87</td><td>0.03</td></tr>
<tr><td>FVC % predicted</td><td>89.03 +/- 27.07</td><td>78.03 +/- 22.80</td><td>Not significant</td></tr>
<tr><td>BMI kg/m^2</td><td>20.2 +/- 3.5</td><td>18.8 +/- 2.7</td><td>Not significant</td></tr>
</table>
<u>Reviewer notes</u>
• Generally indicating that both genotypes are associated with older age at diagnosis and higher rates of pancreatic sufficiency
• Likely representative of these heterozygotes but small samples for comparison
• Not all phenotypic variables clear but assessment should be similar within centres
• No adjustment for confounders aside from age, gender and centre
</td></tr>
<tr><td>MacKenzie et al 2017[23]</td><td>Retrospective cohort
CF Canada Data Registry.
Patients who attended CF clinics 1996-2011 and with the P67L variant.
F508del homozygotes</td><td>N=26 P67L heterozygotes (n=20 F508del/P67L) compared with n=266 F508del homozygotes</td><td>Specific genotype comparison
P67L/F508del
vs F508del/F508del</td><td><u>Phenotypic comparison</u>
<table>
<tr><td>Mean values</td><td>P67L/F508del (n=26)</td><td>F508del/F508del (n=266)</td><td>P value</td></tr>
<tr><td>Age at diagnosis</td><td>18.23 +/- 14.58</td><td>0.92 +/- 0.13</td><td><0.001</td></tr>
<tr><td>Pancreatic insufficient %</td><td>26.9</td><td>99</td><td><0.001</td></tr>
<tr><td>FEV1/FVC annual decline*</td><td colspan="2">Similar pattern of decline in both groups and for both birth cohorts</td><td>Not reported</td></tr>
<tr><td>P. Aeruginosa*</td><td colspan="2">Different patterns and peaks for both groups and both cohorts but no clear difference in colonisation reported</td><td>Not reported</td></tr>
</table>
</td></tr>
</table>

<table>
<tr><th>Study reference</th><th>Design and setting</th><th>Population characteristics</th><th>Genetic comparison</th><th>Outcomes</th></tr>
<tr><td></td><td>from a single clinic in Atlantic Canada

P67L variant not identified through newborn screening; unclear for F508del homozygotes</td><td></td><td></td><td>
<table>
<tr><td>Nutritional status*</td><td colspan="2">Indication of better BMI with increased age with P67L for those born <1965 only</td><td>Not reported</td></tr>
<tr><td>Hospitalisations over 5 years*</td><td>3.63 +/- 0.78</td><td>3.41 +/- 1.00</td><td>>0.05</td></tr>
<tr><td>Hospital days*</td><td>9.62 +/- 4.65</td><td>23.2 +/- 8.00</td><td>0.005</td></tr>
</table>
*longitudinal analyses reported only for birth cohorts with sufficient members: <1965 (n=12 heterozygotes and n=8 homozygotes) and 1981-93 (n=7 heterozygotes and n=107 homozygotes)
Pancreatic status assessed by enzyme replacement therapy

<u>Reviewer notes</u>
• Likely representative of P67L heterozygotes across Canada but still small samples for comparison
• Longitudinal analyses for lung function/infection and nutritional status less likely to be reliable as much smaller samples
• Homozygotes only from single centre, doesn’t report them being age- and sex-matched so unclear how comparable they are to heterozygotes from across Canada
• No adjustment for confounders
</td></tr>
</table>

Table 28. Non-prioritised studies on genotype-phenotype association

Study	Setting	Population	Comparison	Outcomes	Association found?	Finding
Severe vs mild class variants						
Ahmed et al 2003[51]	Toronto clinics 1990-97	633	Severe vs mild class	Pancreatic insufficiency	Yes	Severe 96% vs 2% (no statistical analysis)
Sebro et al 2012[52]	US single centre Dates unclear	435	Severe vs mild class	FEV1	No	Multivariate analysis p=0.98 (values not given)
				P. aeruginosa colonisation	Yes	Severe 66% vs 27% (p<0.001)
				Pancreatic insufficiency	Yes	Severe 94% vs 36% (p<0.001)
Dray et al 2005[53]	France single centre 1997-99	147	Severe vs mild class	Nutrition status: severe malnutrition vs mild/moderate vs well nourished	Yes	Severe: 27% severe malnourished, 29% mild/moderate, 44% well nourished Mild: 8% severe malnourished, 25% mild/moderate, 68% well nourished (p trend <0.01)
F508del homozygotes vs F508del heterozygotes						
Kerem et al 1990[40]	Toronto single centre Dates unclear	293	F508del/F508del vs F508del/other vs other/other	Age at diagnosis	Yes	Homozygotes mean 1.8 yrs vs F508del/other 4.4 yrs vs other/other 8.4 yrs (p<0.001)
				Pancreatic insufficiency	Yes	Homozygotes 98% vs F508del/other 72% vs other/other 36% (p<0.001)

Johansen et al 1991[39]	Denmark single centre 1989	235	F508del/F508del vs F508del/other	Age at diagnosis < 6 months (proportion)	Yes	Homozygotes 94% vs F508del/other 72%, p<0.0005
				FEV1 <70% (proportion)	No	Homozygotes 50% vs F508del/other 49%
				P. aeruginosa colonisation	No	Homozygotes 63% vs F508del/other 54%
Corey et al 1997[54]	Toronto clinics Patients born 1960-74 (surviving >15 yrs with repeat FEV1)	197	F508del/F508del vs F508del/other vs other/other	FEV1 decline	Yes	Mixed model regression: Heterozygotes significantly less decline, lowest for those without F508del (p=0.005 and for slope p=0.048; no difference for intercept aged 5)
Dray et al 2005[53]	France single centre 1997-99	161	F508del/F508del vs F508del/other vs other/other	Nutrition status: severe malnutrition vs mild/moderate vs well nourished	Yes	Homozygotes: 29% severe, malnourished, 32% mild/moderate, 39% well nourished F508del/other: 14% severe, malnourished, 29% mild/moderate, 57% well nourished other/other: 24% severe malnourished, 10% mild/moderate, 66% well nourished (p trend =0.02)
Courtney et al 2007[55]	2 Irish centres 1995-2005	150	F508del/F508del vs F508del/other vs other/other	Survival	No	No difference in genotype proportions of those who died during follow-up vs survived

Gan et al 1995[38]	Netherlands single centre 1995	136	F508del/F508del vs F508del/other vs other/other	Age at diagnosis: proportion diagnosed in adulthood	Yes	Homozygotes 0% vs F508del/other 32% vs other/other 39% (as proportion of genotype in sample; no statistical analysis)
Lester et al 1994[41]	US 3 centres 1990-91	119	F508del/F508del vs F508del/other vs other/other	Age at diagnosis	Yes	Homozygotes mean 1.7 yrs vs F508del/other 3.7yrs vs other/other 4.0 yrs (p<0.05)
				FEV1	No	Homozygotes mean 73% vs F508del/other 66% vs other/other 57%
				Pancreatic insufficiency	No	Homozygotes 93% vs F508del/other 90% vs other/other 79%
				Nutrition status (weight/height %)	No	Homozygotes 96% vs F508del/other 96% vs other/other 94%
Borgo et al 1990[56]	Italy single centre Date unclear	118	F508del/F508del vs F508del/other vs other/other	Pancreatic insufficiency	Yes	Homozygotes 100% vs F508del/other 59% vs other/other 50% (no statistical comparison of proportions but p=0.015 for overall frequency of F508del among analysed chromosomes)
Borgo et al 1993[57]	Italy single centre Date unclear	108	F508del/F508del vs F508del/other vs other/other	Age at diagnosis	No	Homozygotes mean 12mnths vs F508del/other 19mnths vs other/other 15 mnths
				FEV1	No	Homozygotes mean 74% vs F508del/other 76% vs other/other 80%

				P. aeruginosa	Yes	Homozygotes 0.6 vs F508del/other 0.3 vs other/other 0.4 (p=0.005) (colonisation score: 1=present in all samples over 6-8 months)
				Nutrition status (BMI z score)	No	Homozygotes -0.9 vs F508del/other -0.1 vs other/other -0.3
By severe class (1 vs 2)						
Sanders et al 2014[58]	Follow-up of US NBS trial participants (enrolment 1985-94)	132	F508del/F508del vs F508del/other severe class	FEV1	No	4.95 difference in multivariate model (p=0.08) (adjusted for age, BMI, *P. aeruginosa*, recent hospitalisation, meconium ileus)
Geborek and Hjelte 2011[59]	All Swedish patients of Scandinavian prevalence study	266	Within severe class 1/1 (n=18) vs 1/2 (n=78) vs 2/2 (n=170)	FEV1	Yes	-13% difference for class 1/1 vs 1/2 or 2/2 in multivariate model (p=0.01) No difference 1/2 vs 2/2 (adjusted for age, sex, age at diagnosis, BMI, *P. aeruginosa*, diabetes)
By genotype						
Kristidis et al 1992[60]	Toronto single centre Dates unclear	394	By genotype (n=279 F508del/F508del, n=115 F508del/other)	Pancreatic insufficiency	Yes	Homozygotes 99% insufficient Heterozygotes with >5 people: G551D, G542X, 621+1G>T, I507del, N1303K, R560T, 1717-1G>A – all insufficient R117H – all sufficient

Criterion 4 and 8: question 3: screening test accuracy

Table 29. Post-2000 antenatal screening pilot

Study	Design	Screening test	Variants tested	Uptake	Carriers	Outcome
Massie et al 2009[42]	Population-based antenatal screening cohort, Victoria, Australia 2006-08	Pay-for test (Aus $200) offered to women or couples attending a GP: • Prior to pregnancy • In the first 14 weeks of pregnancy Couples screening recommended but mostly stepwise. Method: check swab. Patients provided with information on CF and collection procedure, swab and pre-paid envelope.	12 variant panel known to cover 83.5% of carriers in the general population of the region and 95% of the Ashkenazi Jewish population: • F508del • G551D • G542X • R553X • N1303K • R560T • I507del • W1282X • V520F • 1585-1G→A • 489+1G→T • 3718-2477C→T	Total 3200 screened: • 3000 women • 200 men 100 were couples (200 individuals)	106 carriers detected: • 92 women • 14 men None part of couples screening: • 106 partners tested • 9 carrier couples identified: 3 pre-conception and 6 pregnant	6/6 pregnant couples accepted CVS: 2/6 affected fetuses (PPV 33%) Both terminated. No follow-up of screen negatives.

Appendix 4 — Appraisal for quality and risk of bias

QUIPS quality assessment of genotype-phenotype association studies

Key listed characteristics: Age, gender, ethnicity, country (if applicable), genotype, age and method of diagnosis, baseline characteristics
Key listed confounders: Age, gender, ethnicity, country (if applicable), age and method of diagnosis, treatment (or year of birth/entry to cohort as proxy)

Only information contained within the publication has been considered. Information has not been verified using additional sources, such as accessing national registry data.

QUIPS table adapted from Cochrane Methods Prognosis: Review Tools

Table 30.1

Author and year of publication	**McKone et al 2006**			
Biases	**Issues to consider for judging overall rating of "Risk of bias"**	**Study Methods & Comments**	**Rating of reporting: yes, partial, no, unsure**	**Overall rating of "Risk of bias" for domain: high, moderate, low**

1. Study Participation		**Goal: To judge the risk of selection bias: the likelihood that relationship between prognostic factor (PF) and outcome is different for participants and eligible non-participants. For registry studies, this section has considered the basis for selection for the registry (as all people with CF in the region would theoretically be eligible for analysis) as the equivalent of "participant selection" in a prospective study**		
a)	Source of target population	The source population or population of interest is adequately described for key characteristics (LIST).	Reports characteristics of those genotyped and included in the study but does not give characteristics of the full registry population. Also unclear how representative the registry is of all people with CF in the US	No
b)	Method used to identify population	The sampling frame and recruitment are adequately described, including methods to identify the sample sufficient to limit potential bias (for example, referral patterns in health care)	Eligibility for inclusion in this study is given, but unclear entry into the registry for the CF population	Partial
c)	Recruitment period	Period of recruitment is adequately described	Study observation period for the registry is given (1993-2002) but unclear entry into the registry	Partial
d)	Place of recruitment	Place of recruitment (setting and geographic location) are adequately described	Registry setting and location is given (US) but unclear which clinics or geographical regions this covers	Partial
e)	Inclusion	Inclusion and exclusion	Study includes participants genotyped and	Partial

	and exclusion criteria	criteria are adequately described (including explicit diagnostic criteria)	classified with one of 21 variants of known functional class and frequency >0.1%. As above unclear entry to the registry.		
f)	Adequate study participation	There is adequate participation in the study by eligible individuals (>70%)	Only representative of around 50% in registry due to lack of genotyping or follow-up. Unclear how representative the registry is of all people with CF.	No	
g)	Baseline characteristics	The baseline study sample (individuals entering the study) is adequately described for key characteristics (LIST).	Age and year of entry to registry, gender, proportion mild/severe genotype and baseline characteristics. Age at entry was younger for severe genotypes and follow-up was longer	Partial	
Summary Study participation		**The study sample represents the population of interest on key characteristics, sufficient to limit potential bias for the observed relationship between the PF and outcome.**			**Moderate: many fields uncertain but no clear indication of participation bias in registry**
2. Study Attrition		**Goal: To judge the risk of attrition bias (likelihood that relationship between PF and the outcome are different for completing and non-completing participants). For registry studies, we considered that this section should consider loss of participants from the analysis due to lack of available data, for example.**			

a)	Proportion of baseline sample available for analysis	Response rate is adequate (proportion of study sample completing the study and providing outcome data).	There was no apparent loss of participants among those with genotyping/classification data and with >1 year follow-up (who formed the baseline population for study). But as above they represent only 50% of the potential eligible registry cohort.	No
b)	Attempts to collect information on participants who dropped out	Attempts to collect information on participants who dropped out of the study are described.	No coverage of those who were not genotyped/classified	No
c)	Reasons and potential impact of subjects lost to follow-up	Reasons for loss to follow-up are provided.	As above it's clear that the study has only included those genotyped/classified and with sufficient follow-up but there is no further detail on this.	Partial
d)	Outcome and prognostic factor information on those lost to	Participants lost to follow-up are adequately described for key characteristics (LIST) with no important differences from participants.	Characteristics for those non-genotyped/classified not reported and with inadequate follow-up data not reported.	No

	follow-up			
Study Attrition Summary		**Loss to follow-up (from baseline sample to study population analysed) is not associated with key characteristics (that is, the study data adequately represent the sample) sufficient to limit potential bias to the observed relationship between PF and outcome.**		**High**
3. Prognostic Factor Measurement		**Goal: To judge the risk of measurement bias related to how the PF was measured (differential measurement of PF related to the level of outcome). For studies comparing variant classes this includes whether the system used to classify was adequately described.**		
a)	Definition of the PF	A clear definition or description of 'PF' is provided (including dose, duration of exposure, and clear specification of the measurement method).	Clearly describes studied variants and classification 1-5 using established system	Yes
b)	Valid and Reliable Measurement of PF	Method of PF measurement is valid and reliable to limit misclassification bias (may include relevant outside sources of information on measurement properties, such as blind measurement and limited reliance on	Technical method of genotyping is not given but this is the most established classification system	Partial

		recall).			
c)	Method and Setting of PF Measurement	The method and setting of measurement of PF is the same for all study participants.	Unclear how genotyping was performed across centres and it's likely to have been carried out at different facilities.	Unsure	
d)	Proportion of data on PF available for analysis	Adequate proportion (>70%) of the study sample has complete data for PF variable.	Only 55% of the available subjects with adequate follow-up were genotyped or classified	No	
e)	Method used for missing data	Appropriate methods of imputation are used for missing PF data.	Unclear if any imputation used for genotype data recorded in the registry.	Unsure	
PF Measurement Summary		**PF is adequately measured in study participants to sufficiently limit potential bias.**			**Moderate: due to lack of genotyping and potential variation in lab methods, but expected to be as optimal as**

					possible from registry studies
4. Outcome Measurement		**Goal: To judge the risk of bias related to the measurement of outcome (differential measurement of outcome related to the baseline level of PF).**			
a)	Definition of the Outcome	A clear definition of outcome is provided, including duration of follow-up.	The study is looking at all-cause mortality during the assessment period. It also looks at survival to set age cut-offs but no further detail on definition is given.	Yes	
b)	Valid and Reliable Measurement of Outcome	The method of outcome measurement used is adequately valid and reliable to limit misclassification bias (may include relevant outside sources of information on measurement properties, also characteristics, such as blind measurement and confirmation of outcome with valid and reliable test).	Unclear how deaths were identified. Also mortality analyses report including “patients who died after transplant” but it’s not explicit whether need for transplant itself has been considered as mortality	Unsure	

c)	Method and Setting of Outcome Measurement	The method and setting of outcome measurement is the same for all study participants.	Unclear how mortality was identified, though it's expected the same method may have been used for all participants.	Unsure	
Outcome Measurement Summary		**Outcome of interest is adequately measured in study participants to sufficiently limit potential bias.**			**Moderate: as the outcome is mortality any error may be expected to be consistent across participants**
5. Study Confounding		**Goal: To judge the risk of bias due to confounding (where the effect of the PF is distorted by another factor that is related to both the PF and outcome).**			
a)	Important Confounders Measured	Important confounders including treatments are measured (key LIST variables)	Adjusts for phenotypic variables, year of entry to cohort and centre population size – the latter are assumed as proxies for care received (which is the optimal any study gets for adjustment for treatment). Age, gender, ethnicity and type of presentation not assessed	Partial	

b)	Definition of the confounding factor	Clear definitions of the important confounders measured are provided (including dose, level, and duration of exposure).	Explains how phenotypic variables were measured and that they were collected during year of entry to the study. Though as below it's uncertain how reliable these may be. There is no detail for example on dose of ERT, duration of use (just presence yes or no) but this is as standard. No detail on information about centre size/year of entry	Partial
c)	Valid and Reliable Measurement of Confounders	Measurement of all important confounders is adequately valid and reliable (may include relevant outside sources of information on measurement properties, also characteristics, such as blind measurement and limited reliance on recall).	As above related to phenotypic variables it's unclear how reliably they reflect the variables (such as ERT use for sufficiency, positive *P. aeruginosa* culture in past year). Unclear how other centre variables and age of entry to cohort was assessed.	Unsure
d)	Method and Setting of Confounding Measureme	The method and setting of confounding measurement are the same for all study participants.	Multicentre registry study and so likely variability in how measured across centres and how they may have been entered into registry	No

	nt			
e)	Method used for missing data	Appropriate methods are used if imputation is used for missing confounder data.	Unsure whether there may have been missing data on confounders or how this was managed.	Unsure
f)	Appropriate Accounting for Confoundin g	Important potential confounders are accounted for in the study design (for example, matching for key variables, stratification, or initial assembly of comparable groups).	No matching or stratification	No
		Important potential confounders are accounted for in the analysis (that is, appropriate adjustment).	As above some relevant confounders are adjusted for	Partial
Study Confounding Summary		**Important potential confounders are appropriately accounted for, limiting potential bias with respect to the relationship between PF and outcome.**		**Moderate: on the basis that as a registry study this has attempted to**

					adjust for some relevant confounders
6. Statistical Analysis and Reporting		**Goal: To judge the risk of bias related to the statistical analysis and presentation of results.**			
a)	Presentation of analytical strategy	There is sufficient presentation of data to assess the adequacy of the analysis.	Statistical methods described	Yes	
b)	Model development strategy	The strategy for model building (inclusion of variables in the statistical model) is appropriate and based on a conceptual framework or model. The selected statistical model is adequate for the design of the study.	Builds Cox proportional hazards model to assess genotype as predictor of mortality	Yes	
c)	Reporting of results	There is no selective reporting of results.	None apparent	No	
Statistical Analysis and		**The statistical analysis is appropriate for the design of the study, limiting potential for presentation of invalid or spurious results.**			**Low**

Presentation Summary	
Summary McKone et al 2006: Participation moderate; Attrition high; PF moderate; Outcome moderate; Confounding moderate; Statistical Analysis low	

Table 30.2

Author and year of publication	**McKone et al 2003**			
Biases	**Issues to consider for judging overall rating of "Risk of bias"**	**Study Methods & Comments**	**Rating of reporting: yes, partial, no, unsure**	**Overall rating of "Risk of bias" for domain: high, moderate, low**
1. Study Participation	**Goal: To judge the risk of selection bias: the likelihood that relationship between prognostic factor (PF) and outcome is different for participants and eligible non-participants. For registry studies, this section has considered the basis for selection for the registry (as all people with CF in the region would theoretically be eligible for analysis) as the equivalent of "participant selection" in a prospective study**			
a) Source of target population	The source population or population of interest is adequately described for key	Reports characteristics of the total registry cohort, those genotyped and included in the study. Explains that the CF registry has collected demographic and clinical data	Yes	

		characteristics (LIST).	since 1964 and covers over 85% of people from across the country.	
b)	Method used to identify population	The sampling frame and recruitment are adequately described, including methods to identify the sample sufficient to limit potential bias (for example, referral patterns in health care)	Includes those who were genotyped at any time within the follow-up period. Specific process by which patients are entered into the registry is unclear and unclear whether certain geographic regions may have limited clinic coverage.	Partial
c)	Recruitment period	Period of recruitment is adequately described	Study observation period for the registry is given (1991-1999) but unclear specifically how patients are entered into the registry	Partial
d)	Place of recruitment	Place of recruitment (setting and geographic location) are adequately described	Says that this covers CF accredited centres throughout the US and covers 85% of those with CF across the country.	Yes
e)	Inclusion and exclusion criteria	Inclusion and exclusion criteria are adequately described (including explicit diagnostic criteria)	Study includes participants genotyped and classified with 24 variants present in over 84% of those with CF. The registry covers 85% but unclear whether there may be less clinic coverage in certain geographic regions accounting for	Partial

			those not entered into the registry		
f)	Adequate study participation	There is adequate participation in the study by eligible individuals (>70%)	The registry is representative of 85% of those in the US so should give coverage but only 62% of the available cohort are genotyped.	Partial	
g)	Baseline characteristics	The baseline study sample (individuals entering the study) is adequately described for key characteristics (LIST).	Age, gender, ethnicity, age at diagnosis, baseline characteristics and number of deaths are given and can be compared for the full registry and those genotyped.	Yes	
Summary Study participation		**The study sample represents the population of interest on key characteristics, sufficient to limit potential bias for the observed relationship between the PF and outcome.**			**Low: better indication in the study of how representative the database is and full description baseline population**
2. Study Attrition		**Goal: To judge the risk of attrition bias (likelihood that relationship between PF and the outcome are different for completing and non-completing participants). For registry studies, we considered that this section should consider loss of participants from the analysis due to lack of available data, for example.**			

a)	Proportion of baseline sample available for analysis	Response rate is adequate (proportion of study sample completing the study and providing outcome data).	There was no apparent loss of participants among those with genotyping/classification data available (who formed the baseline population for study). But as above they represent only 62% of the potential eligible registry cohort.	No
b)	Attempts to collect information on participants who dropped out	Attempts to collect information on participants who dropped out of the study are described.	The study provides characteristics for the full registry cohort and those with genotyping data available.	Yes
c)	Reasons and potential impact of subjects lost to follow-up	Reasons for loss to follow-up are provided.	As above it's clear that the study has only included those genotyped/classified but there is no further detail on why participants may not have been genotyped.	Partial
d)	Outcome and prognostic factor information on those lost to	Participants lost to follow-up are adequately described for key characteristics (LIST) with no important differences from participants.	Characteristics for those non-genotyped/classified are reported. Most differences are only minor except for perhaps mortality (12% of total cohort died compared with 9% of the genotyped cohort). However, no statistical comparison is given so it's unclear if these are significant.	Partial

follow-up				
Study Attrition Summary	**Loss to follow-up (from baseline sample to study population analysed) is not associated with key characteristics (that is, the study data adequately represent the sample) sufficient to limit potential bias to the observed relationship between PF and outcome.**			**Moderate: characteristics have been given for non-genotyped cohort with no obvious differences**
3. Prognostic Factor Measurement	**Goal: To judge the risk of measurement bias related to how the PF was measured (differential measurement of PF related to the level of outcome). For studies comparing variant classes this includes whether the system used to classify was adequately described.**			
a) Definition of the PF	A clear definition or description of 'PF' is provided (including dose, duration of exposure, and clear specification of the measurement method).	Clearly describes studied variants and classification 1-5 using established system.	Yes	
b) Valid and Reliable Measurement of PF	Method of PF measurement is valid and reliable to limit misclassification bias (may include relevant	Technical method of genotyping is not given. This is the most established classification system. G85E has since been reclassified but only constitutes a small sample of people.	Partial	

		outside sources of information on measurement properties, such as blind measurement and limited reliance on recall).			
c)	Method and Setting of PF Measurement	The method and setting of measurement of PF is the same for all study participants.	Unclear how genotyping was performed across centres and it's likely to have been carried out at different facilities.	Unsure	
d)	Proportion of data on PF available for analysis	Adequate proportion (>70%) of the study sample has complete data for PF variable.	Only 65% of the available cohort genotyped or classified.	No	
e)	Method used for missing data	Appropriate methods of imputation are used for missing PF data.	Unclear if any imputation used for genotype data recorded in the registry.	Unsure	
PF Measurement Summary		**PF is adequately measured in study participants to sufficiently limit potential bias.**			**Moderate: due to lack of genotyping and potential variation in lab methods, but expected to be**

					as optimal as possible from registry studies
4. Outcome Measurement		**Goal: To judge the risk of bias related to the measurement of outcome (differential measurement of outcome related to the baseline level of PF).**			
a)	Definition of the Outcome	A clear definition of outcome is provided, including duration of follow-up.	The study is looking at mortality rates, which have been calculated by dividing the number of deaths by the number of person-years at risk. Standardised for age and gender distribution.	Yes	
b)	Valid and Reliable Measurement of Outcome	The method of outcome measurement used is adequately valid and reliable to limit misclassification bias (may include relevant outside sources of information on measurement properties, also characteristics, such as blind measurement and confirmation of	Clearly describes what was considered as mortality, including those who needed transplant. Though it's not explicitly explained how deaths may have been identified within the registry.	Partial	

		outcome with valid and reliable test).			
c)	Method and Setting of Outcome Measurement	The method and setting of outcome measurement is the same for all study participants.	Unclear how mortality was assessed, though it's expected the same method may have been used for all participants.	Unsure	
Outcome Measurement Summary		**Outcome of interest is adequately measured in study participants to sufficiently limit potential bias.**			**Moderate: as the outcome is mortality any error may be expected to be consistent across participants**
5. Study Confounding		**Goal: To judge the risk of bias due to confounding (where the effect of the PF is distorted by another factor that is related to both the PF and outcome).**			
a)	Important Confounders Measured	Important confounders including treatments are measured (key LIST variables)	Age and gender accounted for but otherwise no adjustment for confounders.	Partial	
b)	Definition of the confounding factor	Clear definitions of the important confounders measured are provided (including	Explains that variables are assessed for a mean 15 year old cohort in which 52% of the cohort were male but otherwise no adjustment	Partial	

		dose, level, and duration of exposure).		
c)	Valid and Reliable Measurement of Confounders	Measurement of all important confounders is adequately valid and reliable (may include relevant outside sources of information on measurement properties, also characteristics, such as blind measurement and limited reliance on recall).	Only age and gender assessed	NA
d)	Method and Setting of Confounding Measurement	The method and setting of confounding measurement are the same for all study participants.	Only age and gender assessed	NA
e)	Method used for missing data	Appropriate methods are used if imputation is used for missing confounder data.	Not applicable as not measured	NA

f)	Appropriate Accounting for Confoundin g	Important potential confounders are accounted for in the study design (for example, matching for key variables, stratification, or initial assembly of comparable groups).	Stratification for age and gender but no other confounders adjusted for.	Partial	
		Important potential confounders are accounted for in the analysis (that is, appropriate adjustment).	Limited measured other than age and gender.	No	
Study Confounding Summary		**Important potential confounders are appropriately accounted for, limiting potential bias with respect to the relationship between PF and outcome.**			**High**
6. Statistical Analysis and Reporting		**Goal: To judge the risk of bias related to the statistical analysis and presentation of results.**			
a)	Presentation of analytical strategy	There is sufficient presentation of data to assess the adequacy of the analysis.	Statistical methods described	Yes	

b)	Model development strategy	The strategy for model building (inclusion of variables in the statistical model) is appropriate and based on a conceptual framework or model. The selected statistical model is adequate for the design of the study.	The study has calculated standardised mortality rates and used linear regression to compare variables between groups. P values for significance are given.	Yes	
c)	Reporting of results	There is no selective reporting of results.	None apparent	No	
Statistical Analysis and Presentation Summary		**The statistical analysis is appropriate for the design of the study, limiting potential for presentation of invalid or spurious results.**			**Low**

Summary McKone et al 2003: Participation low; Attrition moderate; PF moderate; Outcome moderate; Confounding high; Statistical Analysis low

Table 30.3

Author and year of publication	**Lai et al 2004**			
Biases	**Issues to**	**Study Methods & Comments**	**Rating of**	**Overall**

		consider for judging overall rating of "Risk of bias"		**reporting: yes, partial, no, unsure**	**rating of "Risk of bias" for domain: high, moderate, low**
1. Study Participation		**Goal: To judge the risk of selection bias: the likelihood that relationship between prognostic factor (PF) and outcome is different for participants and eligible non-participants. For registry studies, this section has considered the basis for selection for the registry (as all people with CF in the region would theoretically be eligible for analysis) as the equivalent of "participant selection" in a prospective study**			
a)	Source of target population	The source population or population of interest is adequately described for key characteristics (LIST).	Reports characteristics for 85% of the registry cohort with >1 follow-up and so available for time to event analysis, but doesn't give data for the full registry cohort or indicate how representative the registry is of all people with CF in the US.	Partial	
b)	Method used to identify population	The sampling frame and recruitment are adequately described, including methods to identify the sample sufficient to limit potential bias (for example, referral patterns in health care)	Eligibility for inclusion in this study is given, but unclear entry into the registry for the CF population	Partial	

c)	Recruitment period	Period of recruitment is adequately described	Study observation period for the registry is given (1986-2000) but unclear entry into the registry	Partial	
d)	Place of recruitment	Place of recruitment (setting and geographic location) are adequately described	Registry setting and location is given (US) but less information on which clinics or geographical regions this covers	Partial	
e)	Inclusion and exclusion criteria	Inclusion and exclusion criteria are adequately described (including explicit diagnostic criteria)	Study includes participants with >1 follow-up and data on method of diagnosis. Genotype analysis includes those genotyped and classified. As above unclear entry to the registry.	Partial	
f)	Adequate study participation	There is adequate participation in the study by eligible individuals (>70%)	Variants specified and classified for only 49% of the included registry population with follow-up data and information on method of diagnosis. Unclear how representative the registry is of all people with CF.	No	
g)	Baseline characteristics	The baseline study sample (individuals entering the study) is adequately described for key characteristics (LIST).	Gender, age and method of diagnosis and genotyped. Doesn't give current age, ethnicity or other phenotypic variables. Presentation reportedly differed for severe genotypes (most of whom were identified by meconium ileus)	Partial	
Summary Study		**The study sample represents the population of interest on key characteristics, sufficient to limit**			**Moderate:**

participation		**potential bias for the observed relationship between the PF and outcome.**			**many fields uncertain but no clear indication of participation bias in registry**
2. Study Attrition		**Goal: To judge the risk of attrition bias (likelihood that relationship between PF and the outcome are different for completing and non-completing participants). For registry studies, we considered that this section should consider loss of participants from the analysis due to lack of available data, for example.**			
a)	Proportion of baseline sample available for analysis	Response rate is adequate (proportion of study sample completing the study and providing outcome data).	Only 49% of the potential eligible registry cohort had classified genotype for analysis with only 25-33% with data for FEV1 and infection analysis	No	
b)	Attempts to collect information on participants who dropped out	Attempts to collect information on participants who dropped out of the study are described.	No coverage of those who were not genotyped/classified	No	
c)	Reasons	Reasons for loss to	The study has only included those	Partial	

	and potential impact of subjects lost to follow-up	follow-up are provided.	genotyped/classified and with sufficient follow-up for the main analysis. For FEV1 and infective colonisation people with FEV1<70% and infected at first visit were excluded.		
d)	Outcome and prognostic factor information on those lost to follow-up	Participants lost to follow-up are adequately described for key characteristics (LIST) with no important differences from participants.	No clear differences in proportions diagnosed by different method , though no statistical analysis and other characteristics not compared. For phenotypic assessment exclusion of those with FEV1<70% and with early infection may exclude more severe genotypes.	No	
Study Attrition Summary		**Loss to follow-up (from baseline sample to study population analysed) is not associated with key characteristics (that is, the study data adequately represent the sample) sufficient to limit potential bias to the observed relationship between PF and outcome.**			**High**
3. Prognostic Factor Measurement		**Goal: To judge the risk of measurement bias related to how the PF was measured (differential measurement of PF related to the level of outcome). For studies comparing variant classes this includes whether the system used to classify was adequately described.**			
a)	Definition of the PF	A clear definition or description of 'PF' is provided (including dose, duration of exposure, and clear specification of the	Lists studied variants and classification 1-5	Yes	

		measurement method).			
b)	Valid and Reliable Measurement of PF	Method of PF measurement is valid and reliable to limit misclassification bias (may include relevant outside sources of information on measurement properties, such as blind measurement and limited reliance on recall).	Technical method of genotyping is not given. Classification system also has several discrepancies with differences in grouping of G85E, 2789+5G>A and A455E between severe and mild classes. Also lists 2184delA, 1898+1G>A and 711+1G>A put has group 1 when other studies have put these are unclassified	No	
c)	Method and Setting of PF Measurement	The method and setting of measurement of PF is the same for all study participants.	Unclear how genotyping was performed across centres and it's likely to have been carried out at different facilities.	Unsure	
d)	Proportion of data on PF available for analysis	Adequate proportion (>70%) of the study sample has complete data for PF variable.	Only 66% of the available subjects with adequate follow-up were genotyped and 49% classified and used in analysis. As above fewer for FEV1 and infection analysis	No	
e)	Method used for missing data	Appropriate methods of imputation are used for missing PF data.	Unclear if any imputation used for genotype data recorded in the registry.	Unsure	
PF		**PF is adequately measured in study participants to sufficiently limit potential bias.**			**High: due to**

Measurement Summary				**lack of genotyping and also discrepancies in classification**
4. Outcome Measurement	**Goal: To judge the risk of bias related to the measurement of outcome (differential measurement of outcome related to the baseline level of PF).**			
a) Definition of the Outcome	A clear definition of outcome is provided, including duration of follow-up.	The study is looking primarily at survival but gives no indication how this is measured and the analysis just looks at the association of "longer" with "shorter". People aged >6 with FEV1 first measure >70% and >1 follow-up needed for assessment of FEV1 and infective colonisation (so includes smaller proportion of genotyped sample).	No	
b) Valid and Reliable Measurement of Outcome	The method of outcome measurement used is adequately valid and reliable to limit misclassification bias (may include relevant outside	Unclear how deaths were identified, whether transplant was considered in mortality and gives no indication of what survival may be other than "longer" or "shorter". Phenotypic variables looking at time to FEV1 measure <70% and to *P.aeruginosa*	No	

		sources of information on measurement properties, also characteristics, such as blind measurement and confirmation of outcome with valid and reliable test).	colonisation, though risk comparison are likely again to be "longer" or "shorter" without definition.		
c)	Method and Setting of Outcome Measurement	The method and setting of outcome measurement is the same for all study participants.	Unclear how mortality was identified (though it's expected the same method may have been used for all participants). Other variables may have differed for outcome assessment.	Unsure	
Outcome Measurement Summary		**Outcome of interest is adequately measured in study participants to sufficiently limit potential bias.**			**High**
5. Study Confounding		**Goal: To judge the risk of bias due to confounding (where the effect of the PF is distorted by another factor that is related to both the PF and outcome).**			
a)	Important Confounders Measured	Important confounders including treatments are measured (key LIST variables)	No apparent adjustment for any confounders in genotype assessment (only considers cohort year when looking at link between survival and diagnostic group)	No	
b)	Definition of the confounding	Clear definitions of the important confounders measured are provided	None measured	No	

	factor	(including dose, level, and duration of exposure).		
c)	Valid and Reliable Measurement of Confounders	Measurement of all important confounders is adequately valid and reliable (may include relevant outside sources of information on measurement properties, also characteristics, such as blind measurement and limited reliance on recall).	None measured	No
d)	Method and Setting of Confounding Measurement	The method and setting of confounding measurement are the same for all study participants.	None measured	No
e)	Method used for missing data	Appropriate methods are used if imputation is used for missing confounder data.	None measured	No

f)	Appropriate Accounting for Confoundin g	Important potential confounders are accounted for in the study design (for example, matching for key variables, stratification, or initial assembly of comparable groups).	No matching or stratification	No	
		Important potential confounders are accounted for in the analysis (that is, appropriate adjustment).	None measured	No	
Study Confounding Summary		**Important potential confounders are appropriately accounted for, limiting potential bias with respect to the relationship between PF and outcome.**			**High**
6. Statistical Analysis and Reporting		**Goal: To judge the risk of bias related to the statistical analysis and presentation of results.**			
a)	Presentation of analytical strategy	There is sufficient presentation of data to assess the adequacy of the analysis.	Statistical methods described, though as below the primary analysis is looking at link between diagnostic group and survival and lung function outcomes	Partial	

b)	Model development strategy	The strategy for model building (inclusion of variables in the statistical model) is appropriate and based on a conceptual framework or model. The selected statistical model is adequate for the design of the study.	Builds Cox proportional hazards model to assess effects of baseline risk factors on survival and lung function. Genotype assessment was only a secondary analysis.	Partial	
c)	Reporting of results	There is no selective reporting of results.	None apparent	No	
Statistical Analysis and Presentation Summary		**The statistical analysis is appropriate for the design of the study, limiting potential for presentation of invalid or spurious results.**			**Moderate: not the primary design of the analysis to look at effect of genotype**

Summary Lai et al 2004: Participation moderate; Attrition high; PF high; Outcome high; Confounding high; Statistical Analysis moderate

Table 30.4

Author and year of publication	**O'Connor et al 2002**			
Biases	**Issues to**	**Study Methods & Comments**	**Rating of**	**Overall**

		consider for judging overall rating of "Risk of bias"		reporting: yes, partial, no, unsure	rating of "Risk of bias" for domain: high, moderate, low
1. Study Participation		**Goal: To judge the risk of selection bias: the likelihood that relationship between prognostic factor (PF) and outcome is different for participants and eligible non-participants. For registry studies, this section has considered the basis for selection for the registry (as all people with CF in the region would theoretically be eligible for analysis) as the equivalent of "participant selection" in a prospective study**			
a)	Source of target population	The source population or population of interest is adequately described for key characteristics (LIST).	Reports characteristics of the total registry cohort for age gender, ethnicity, age at diagnosis and method of presentation, proportion genotyped and socioeconomic status. Study describes that the registry has maintained annual information on all patients seen at CF Care Centres since 1982 but unclear whether some patients could be missed.	Partial	
b)	Method used to identify population	The sampling frame and recruitment are adequately described, including methods to identify the sample	The study included those who were genotyped during the follow-up period and with data on all outcomes. Specific process by which patients are entered into the registry is unclear and	Partial	

		sufficient to limit potential bias (for example, referral patterns in health care)	unclear whether certain geographic regions may have limited clinic coverage.	
c)	Recruitment period	Period of recruitment is adequately described	Study observation period for the registry is given (1982-1998) but unclear specifically how patients are entered into the registry	Partial
d)	Place of recruitment	Place of recruitment (setting and geographic location) are adequately described	Says that this covers CF accredited centres throughout the US though as above unclear whether all regions could be covered.	Partial
e)	Inclusion and exclusion criteria	Inclusion and exclusion criteria are adequately described (including explicit diagnostic criteria)	Study includes participants genotyped and with other data on other variables. Unclear whether there may be less clinic coverage in certain geographic regions.	Partial
f)	Adequate study participation	There is adequate participation in the study by eligible individuals (>70%)	Only 50% of registry population had genotyping data and unclear how representative the study is. Unclear how representative the registry is.	No
g)	Baseline characteristics	The baseline study sample (individuals entering the study) is adequately described	Age, gender, ethnicity, age and method of diagnosis, genotyping and number of deaths. No phenotypic variables or full comparison variables for those genotyped.	Partial

		for key characteristics (LIST).			
Summary Study participation		**The study sample represents the population of interest on key characteristics, sufficient to limit potential bias for the observed relationship between the PF and outcome.**			**Moderate: many fields uncertain but no clear indication of participation bias in registry**
2. Study Attrition		**Goal: To judge the risk of attrition bias (likelihood that relationship between PF and the outcome are different for completing and non-completing participants). For registry studies, we considered that this section should consider loss of participants from the analysis due to lack of available data, for example.**			
a)	Proportion of baseline sample available for analysis	Response rate is adequate (proportion of study sample completing the study and providing outcome data).	There was no apparent loss of participants among those with genotyping/classification data available (who formed the baseline population for study). But as above they represent only 50% of the potential eligible registry cohort.	No	
b)	Attempts to collect information on participants who	Attempts to collect information on participants who dropped out of the study are described.	The study provides characteristics for the full registry cohort. However, as below does not give separate comparative data on characteristics of those who were genotyped.	Partial	

	dropped out				
c)	Reasons and potential impact of subjects lost to follow-up	Reasons for loss to follow-up are provided.	As above it's clear that the study has only included those genotyped/classified but there is no further detail on why participants may not have been genotyped.	Partial	
d)	Outcome and prognostic factor information on those lost to follow-up	Participants lost to follow-up are adequately described for key characteristics (LIST) with no important differences from participants.	Does not give comparison data for those genotyped so unclear whether there may be differences.	No	
Study Attrition Summary		**Loss to follow-up (from baseline sample to study population analysed) is not associated with key characteristics (that is, the study data adequately represent the sample) sufficient to limit potential bias to the observed relationship between PF and outcome.**			**High: characteristics cannot be compared between genotyped/non-genotyped cohorts**
3. Prognostic Factor		**Goal: To judge the risk of measurement bias related to how the PF was measured (differential measurement of PF related to the level of outcome). For studies comparing variant classes this**			

Measurement		**includes whether the system used to classify was adequately described.**		
a)	Definition of the PF	A clear definition or description of 'PF' is provided (including dose, duration of exposure, and clear specification of the measurement method).	Only comparison is F508del homozygotes and heterozygotes. Unknown second variant.	Partial
b)	Valid and Reliable Measurement of PF	Method of PF measurement is valid and reliable to limit misclassification bias (may include relevant outside sources of information on measurement properties, such as blind measurement and limited reliance on recall).	No description is given on how genotype has been assessed. Technical method of genotyping is not given.	No
c)	Method and Setting of PF Measurement	The method and setting of measurement of PF is the same for all study participants.	Unclear how genotyping was performed across centres and it's likely to have been carried out at different facilities.	Unsure

d)	Proportion of data on PF available for analysis	Adequate proportion (>70%) of the study sample has complete data for PF variable.	Only 50% of the available cohort genotyped or classified.	No	
e)	Method used for missing data	Appropriate methods of imputation are used for missing PF data.	Unclear if any imputation used for genotype data recorded in the registry.	Unsure	
PF Measurement Summary		**PF is adequately measured in study participants to sufficiently limit potential bias.**			**High: limited information on genotyping and comparison is less informative**
4. Outcome Measurement		**Goal: To judge the risk of bias related to the measurement of outcome (differential measurement of outcome related to the baseline level of PF).**			
a)	Definition of the Outcome	A clear definition of outcome is provided, including duration of follow-up.	The study is looking at survival and gives the total number of deaths during the years of follow-up but gives no further information.	Partial	
b)	Valid and Reliable Measurement of Outcome	The method of outcome measurement used is adequately valid and reliable to limit	Limited definition of the outcome. Does not mention whether transplant was considered. Risk analyses just looks at whether genotypes had comparatively increased or decreased risk of death but limited	No	

		misclassification bias (may include relevant outside sources of information on measurement properties, also characteristics, such as blind measurement and confirmation of outcome with valid and reliable test).	interpretation could be made from this (e.g. age of death). It's not explained how deaths may have been identified within the registry.		
c)	Method and Setting of Outcome Measurement	The method and setting of outcome measurement is the same for all study participants.	Unclear how mortality was assessed, though it's expected the same method may have been used for all participants.	Unsure	
Outcome Measurement Summary		**Outcome of interest is adequately measured in study participants to sufficiently limit potential bias.**			**Moderate: as the outcome is mortality any error in measurement is expected to be consistent across participants,**

					but limited interpretation can be made from the results
5. Study Confounding		**Goal: To judge the risk of bias due to confounding (where the effect of the PF is distorted by another factor that is related to both the PF and outcome).**			
a)	Important Confounders Measured	Important confounders including treatments are measured (key LIST variables)	Age and method of diagnosis, gender, ethnicity, socioeconomic factors, but no analysis of treatment.	Partial	
b)	Definition of the confounding factor	Clear definitions of the important confounders measured are provided (including dose, level, and duration of exposure).	Variables that have been assessed are clearly described.	Yes	
c)	Valid and Reliable Measurement of Confounders	Measurement of all important confounders is adequately valid and reliable (may include relevant outside	There may be some inaccuracies in estimating household income from postcode though this was not set as one of the key variables. Unsure whether there may have been any inaccuracies in the registry data for other variables.	Unsure	

		sources of information on measurement properties, also characteristics, such as blind measurement and limited reliance on recall).		
d)	**Method and Setting of Confounding Measurement**	The method and setting of confounding measurement are the same for all study participants.	Multicentre registry study and so likely variability in how measured across centres and how they may have been entered into registry	No
e)	**Method used for missing data**	Appropriate methods are used if imputation is used for missing confounder data.	Only patients with complete socioeconomic data were assessed but unclear whether any imputation may have been used for missing data.	Unsure

f)	**Appropriate Accounting for Confounding**	Important potential confounders are accounted for in the study design (for example, matching for key variables, stratification, or initial assembly of comparable groups).	Matching or stratification not performed.	No	
		Important potential confounders are accounted for in the analysis (that is, appropriate adjustment).	Multivariate analysis for the above factors.	Yes	
Study Confounding Summary		**Important potential confounders are appropriately accounted for, limiting potential bias with respect to the relationship between PF and outcome.**			**Moderate**
6. Statistical Analysis and Reporting		**Goal: To judge the risk of bias related to the statistical analysis and presentation of results.**			
a)	Presentation of analytical strategy	There is sufficient presentation of data to assess the adequacy of the	Statistical methods described	Yes	

		analysis.			
b)	Model development strategy	The strategy for model building (inclusion of variables in the statistical model) is appropriate and based on a conceptual framework or model. The selected statistical model is adequate for the design of the study.	Multivariate analysis performed to predict survival and identify the case-mix to adjust for in analysis of mortality in CF. Kaplan-Meier survival analysis for the effect of each (confounding) variable on survival time. Cox proportional hazard regression used to conduct multivariate tests of the significance of each variable.	Yes	
c)	Reporting of results	There is no selective reporting of results.	None apparent	No	
Statistical Analysis and Presentation Summary		The statistical analysis is appropriate for the design of the study, limiting potential for presentation of invalid or spurious results.			Low

Summary O'Connor et al 2002: Participation moderate; Attrition high; PF high; Outcome moderate; Confounding moderate; Statistical Analysis low

Table 30.5

Author and year of publication	Simmonds et al 2009			
Biases	**Issues to consider for judging overall rating of "Risk of bias"**	**Study Methods & Comments**	**Rating of reporting: yes, partial, no, unsure**	**Overall rating of "Risk of bias" for domain: high, moderate, low**
1. Study Participation	**Goal: To judge the risk of selection bias: the likelihood that relationship between prognostic factor (PF) and outcome is different for participants and eligible non-participants. For registry studies, this section has considered the basis for selection for the registry (as all people with CF in the region would theoretically be eligible for analysis) as the equivalent of "participant selection" in a prospective study**			
a) Source of target population	The source population or population of interest is adequately described for key characteristics (LIST).	Age at diagnosis, gender, phenotypic variables, genotype and median age of death. Incomplete comparison data for full adult registry population aged >16 or specifically aged >35 years. Full registry population aged >40 not described. Also unclear from this paper how representative the registry is of all adults with CF in the UK.	Partial	

b)	Method used to identify population	The sampling frame and recruitment are adequately described, including methods to identify the sample sufficient to limit potential bias (for example, referral patterns in health care)	The study described identifying all adults who had reached >40 years of age on their single centre database in 2004. The study states that all people seen at this centre have been entered into the database since 1965, so this is likely to give coverage of older people within this centre. However, it's unclear how representative they may be of all adults aged >40 in the UK.	Unsure
c)	**Recruitment period**	Period of recruitment is adequately described	As above observation period is given for this centre but unclear how comparison data may have been entered into the fully registry.	Partial
d)	**Place of recruitment**	Place of recruitment (setting and geographic location) are adequately described	Registry setting and location is clearly given for the single centre but less clear for the full registry	Partial
e)	**Inclusion and exclusion criteria**	Inclusion and exclusion criteria are adequately described (including explicit diagnostic criteria)	Study includes participants aged >40 years without transplant and alive to 2004 at their single centre. As above unclear entry to the registry.	Partial
f)	**Adequate study**	There is adequate participation in the	Includes 83% of adults >40 with genotyping data at single centre, but no data is given on	Unsure

	participation	study by eligible individuals (>70%)	the coverage of those in the full registry aged >16, or again how representative they may be of all with CF in the UK.		
g)	**Baseline characteristics**	The baseline study sample (individuals entering the study) is adequately described for key characteristics (LIST).	As above age at diagnosis, gender, phenotypic variables, genotype and median age of death are given for those >40. Incomplete comparison variables for full adult registry population aged >16 or specifically aged >35 years (some imputation of US data instead). Full registry population aged >40 not described. Fewer adults in full registry were diagnosed >16 vs sample >40 years (12 vs 32%).	No	
Summary Study participation		**The study sample represents the population of interest on key characteristics, sufficient to limit potential bias for the observed relationship between the PF and outcome.**			**High: uncertain how representative this single centre may be of all >40 in UK**
2. Study Attrition		**Goal: To judge the risk of attrition bias (likelihood that relationship between PF and the outcome are different for completing and non-completing participants). For registry studies, we considered that this section should consider loss of participants from the analysis due to lack of available data, for**			

		example.		
a)	Proportion of baseline sample available for analysis	Response rate is adequate (proportion of study sample completing the study and providing outcome data).	There was no apparent loss of participants among those with genotyping data aged >40 and they represent 83%. However, coverage of adults in registry cohort for genotype analysis is completely unclear.	No
b)	Attempts to collect information on participants who dropped out	Attempts to collect information on participants who dropped out of the study are described.	Characteristics of all those >40 from single centre given but no comparison to those with genotyping. Incomplete comparison for all adults in registry for all variables, and uncertain how many genotyped.	No
c)	Reasons and potential impact of subjects lost to follow-up	Reasons for loss to follow-up are provided.	The analysis has included only those genotyped in the older cohort but there's no detail for the full registry.	Unsure
d)	Outcome and prognostic factor information on those lost to follow-up	Participants lost to follow-up are adequately described for key characteristics (LIST) with no important differences from participants.	As above no comparison of genotyped/non-genotyped within older cohort or full adult registry.	No

Study Attrition Summary	**Loss to follow-up (from baseline sample to study population analysed) is not associated with key characteristics (that is, the study data adequately represent the sample) sufficient to limit potential bias to the observed relationship between PF and outcome.**			**High**
3. Prognostic Factor Measurement	**Goal: To judge the risk of measurement bias related to how the PF was measured (differential measurement of PF related to the level of outcome). For studies comparing variant classes this includes whether the system used to classify was adequately described.**			
a) Definition of the PF	A clear definition or description of 'PF' is provided (including dose, duration of exposure, and clear specification of the measurement method).	Comparison of F508del homozygotes and heterozygotes. Gives second variant when it is known but the majority were unknown. For the full registry the numbers with the assessed variants are not described.	Partial	
b) Valid and Reliable Measurement of PF	Method of PF measurement is valid and reliable to limit misclassification bias (may include relevant outside sources of information on measurement properties, such as blind measurement and limited reliance on	Technical method of genotyping is not given and there is no detail on how this was identified. No clarity at all for the comparison registry population. Lack of clarity on heterozygotes	No	

		recall).			
c)	Method and Setting of PF Measurement	The method and setting of measurement of PF is the same for all study participants.	Likely to be similar for the single centre though not described. Unclear how genotyping was performed across centres in the registry and it's likely to have been carried out at different facilities.	Unsure	
d)	Proportion of data on PF available for analysis	Adequate proportion (>70%) of the study sample has complete data for PF variable.	83% of the older single centre cohort were genotyped but completely unclear for the comparison population	Partial	
e)	Method used for missing data	Appropriate methods of imputation are used for missing PF data.	Unclear if any imputation used for genotype data recorded in the registry.	Unsure	
PF Measurement Summary		**PF is adequately measured in study participants to sufficiently limit potential bias.**			**High: on the basis of lack of clarity around heterozygotes**
4. Outcome Measurement		**Goal: To judge the risk of bias related to the measurement of outcome (differential measurement of outcome related to the baseline level of PF).**			
a)	Definition of the Outcome	A clear definition of outcome is provided, including duration of follow-up.	The study is looking at survival above set age.	Yes	
b)	Valid and	The method of	As the study is looking at people with	Unsure	

	Reliable Measurement of Outcome	outcome measurement used is adequately valid and reliable to limit misclassification bias (may include relevant outside sources of information on measurement properties, also characteristics, such as blind measurement and confirmation of outcome with valid and reliable test).	specific age there's unlikely to be error, but unclear how regularly data is entered into the full registry and so how up-to-date ages may be.		
c)	Method and Setting of Outcome Measurement	The method and setting of outcome measurement is the same for all study participants.	Likely to be similar for those in the single centre. For the registry, as this is current age unlikely to be affected by UK centre, but unclear how frequently data may be entered from this paper.	Partial	
Outcome Measurement Summary		**Outcome of interest is adequately measured in study participants to sufficiently limit potential bias.**			**Moderate: primarily due to uncertain accuracy on data of current ages**

				within registry
5. Study Confounding		**Goal: To judge the risk of bias due to confounding (where the effect of the PF is distorted by another factor that is related to both the PF and outcome).**		
a)	Important Confounders Measured	Important confounders including treatments are measured (key LIST variables)	None assessed	No
b)	Definition of the confounding factor	Clear definitions of the important confounders measured are provided (including dose, level, and duration of exposure).	None assessed	No
c)	Valid and Reliable Measurement of Confounders	Measurement of all important confounders is adequately valid and reliable (may include relevant outside sources of information on measurement properties, also characteristics, such as blind measurement and limited reliance on	None assessed	No

		recall).			
d)	Method and Setting of Confounding Measurement	The method and setting of confounding measurement are the same for all study participants.	None assessed	No	
e)	Method used for missing data	Appropriate methods are used if imputation is used for missing confounder data.	None assessed	No	
f)	Appropriate Accounting for Confounding	Important potential confounders are accounted for in the study design (for example, matching for key variables, stratification, or initial assembly of comparable groups).	No matching apparent or stratification	No	
		Important potential confounders are accounted for in the analysis (that is, appropriate adjustment).	None assessed	No	
Study		**Important potential confounders are appropriately accounted for, limiting potential bias with respect**			**High**

Confounding Summary	**to the relationship between PF and outcome.**			
6. Statistical Analysis and Reporting	**Goal: To judge the risk of bias related to the statistical analysis and presentation of results.**			
a) Presentation of analytical strategy	There is sufficient presentation of data to assess the adequacy of the analysis.	Fisher's exact test described to be used which seems appropriate for comparison of small groups, but no further detail is given, including no detail on p value for significance	Partial	
b) Model development strategy	The strategy for model building (inclusion of variables in the statistical model) is appropriate and based on a conceptual framework or model. The selected statistical model is adequate for the design of the study.	Does not build a model and gives no further information on the statistical analysis	No	
c) Reporting of results	There is no selective reporting of results.	Predominantly as above unsure how representative those compared are of the full adult registry	Unsure	
Statistical Analysis and	**The statistical analysis is appropriate for the design of the study, limiting potential for presentation of invalid or spurious results.**			**High: small groups for**

Presentation Summary				**comparison and limited data on analysis**
Summary Simmonds et al 2009: Participation high; Attrition high; PF high; Outcome moderate; Confounding high; Statistical Analysis high				

Table 30.6

Author and year of publication	**Badet et al 2004**			
Biases	**Issues to consider for judging overall rating of "Risk of bias"**	**Study Methods & Comments**	**Rating of reporting: yes, partial, no, unsure**	**Overall rating of "Risk of bias" for domain: high, moderate, low**
1. Study Participation	**Goal: To judge the risk of selection bias: the likelihood that relationship between prognostic factor (PF) and outcome is different for participants and eligible non-participants. For registry studies, this section has considered the basis for selection for the registry (as all people with CF in the region would theoretically be eligible for analysis) as the equivalent of "participant selection" in a prospective study**			
a) Source of target	The source population or population of	Gender, age, age and method of diagnosis, genotype and phenotypic variables given for	Partial	

	population	interest is adequately described for key characteristics (LIST).	the older population. No data on ethnicity. Only phenotype and genotype proportions given for the full registry. Registry is said to represent 70% with CF in France, though as below it may not represent all older patients.		
b)	Method used to identify population	The sampling frame and recruitment are adequately described, including methods to identify the sample sufficient to limit potential bias (for example, referral patterns in health care)	Data collected in 1999 for both older cohort and the full registry. Study includes those aged >30 in 1999 and born <1970 but specifies those diagnosed >5 years of age. Therefore could exclude those with milder variants who are diagnosed later. Registry covers 70% of those with CF in France and data is said to be entered every year	No	On basis of excluding older diagnoses which may affect the analysis
c)	Recruitment period	Period of recruitment is adequately described	Study period is given as above and data is entered into the registry yearly.	Yes	
d)	Place of recruitment	Place of recruitment (setting and geographic location) are adequately described	Unclear whether there may be certain geographic locations in France with less coverage.	Unsure	
e)	Inclusion	Inclusion and	Study includes participants aged >30 years	Partial	

	and exclusion criteria	exclusion criteria are adequately described (including explicit diagnostic criteria)	and diagnosed <5 years with comparison to the remainder of the registry. Unclear whether geographic or other factors affect entry into the full registry.		
f)	Adequate study participation	There is adequate participation in the study by eligible individuals (>70%)	Registry covers 70% of those with CF in France. Both variants identified for 82% of older patients and 79% of full registry. As above unclear how representative the study could be of all people >30 (including those diagnosed at later age)	Partial	
g)	Baseline characteristics	The baseline study sample (individuals entering the study) is adequately described for key characteristics (LIST).	As above most variables given for the cohort aged >30 years but incomplete comparison data for those included in the full cohort and those genotyped.	No	
Summary Study participation		**The study sample represents the population of interest on key characteristics, sufficient to limit potential bias for the observed relationship between the PF and outcome.**			**High: predominantly on basis may not represent all those aged >30**
2. Study Attrition		**Goal: To judge the risk of attrition bias (likelihood that relationship between PF and the outcome are different for completing and non-completing participants). For registry studies, we considered that**			

		this section should consider loss of participants from the analysis due to lack of available data, for example.		
a)	Proportion of baseline sample available for analysis	Response rate is adequate (proportion of study sample completing the study and providing outcome data).	As above genotyping was available for sufficient sample size (>70%) in both groups	Yes
b)	Attempts to collect information on participants who dropped out	Attempts to collect information on participants who dropped out of the study are described.	Characteristics of all those >30 given but no comparison to those with genotyping. Incomplete comparison for all adults in registry for all variables, and no distinction for those genotyped/not.	No
c)	Reasons and potential impact of subjects lost to follow-up	Reasons for loss to follow-up are provided.	The analysis has included only those genotyped with no other apparent exclusions.	Yes
d)	Outcome and prognostic factor information	Participants lost to follow-up are adequately described for key characteristics (LIST) with no	As above no comparison of genotyped/non-genotyped within older cohort or full adult registry.	No

	on those lost to follow-up	important differences from participants.		
Study Attrition Summary		**Loss to follow-up (from baseline sample to study population analysed) is not associated with key characteristics (that is, the study data adequately represent the sample) sufficient to limit potential bias to the observed relationship between PF and outcome.**		**Moderate: on basis that sufficient proportion of both groups were genotyped but no comparison data given**
3. Prognostic Factor Measurement		**Goal: To judge the risk of measurement bias related to how the PF was measured (differential measurement of PF related to the level of outcome). For studies comparing variant classes this includes whether the system used to classify was adequately described.**		
a)	Definition of the PF	A clear definition or description of 'PF' is provided (including dose, duration of exposure, and clear specification of the measurement method).	Comparison of F508del homozygotes and heterozygotes. Gives examples of second variant for heterozygotes. For full registry the other variants are unknown.	Partial
b)	Valid and	Method of PF	Technical method of genotyping is not given	No

	Reliable Measurement of PF	measurement is valid and reliable to limit misclassification bias (may include relevant outside sources of information on measurement properties, such as blind measurement and limited reliance on recall).	and there is no detail on how this was identified. Lack of clarity on heterozygotes		
c)	Method and Setting of PF Measurement	The method and setting of measurement of PF is the same for all study participants.	Unclear how genotyping was performed across centres in the registry and it's likely to have been carried out at different facilities.	Unsure	
d)	Proportion of data on PF available for analysis	Adequate proportion (>70%) of the study sample has complete data for PF variable.	As above sufficient sample of both groups genotyped	Yes	
e)	Method used for missing data	Appropriate methods of imputation are used for missing PF data.	Unclear if any imputation used for genotype data recorded in the registry.	Unsure	
PF Measurement Summary		**PF is adequately measured in study participants to sufficiently limit potential bias.**			**High: on the basis of lack of clarity**

					around heterozygotes
4. Outcome Measurement		**Goal: To judge the risk of bias related to the measurement of outcome (differential measurement of outcome related to the baseline level of PF).**			
a)	Definition of the Outcome	A clear definition of outcome is provided, including duration of follow-up.	The study is looking at survival above set age.	Yes	
b)	Valid and Reliable Measurement of Outcome	The method of outcome measurement used is adequately valid and reliable to limit misclassification bias (may include relevant outside sources of information on measurement properties, also characteristics, such as blind measurement and confirmation of outcome with valid and reliable test).	As the study is looking at people with specific age there's unlikely to be error, and data has been entered into the register every year so should be accurate	Yes	
c)	Method and Setting of	The method and setting of outcome	Likely to be unaffected as the outcome is current age.	Partial	

	Outcome Measurement	measurement is the same for all study participants.			
Outcome Measurement Summary		**Outcome of interest is adequately measured in study participants to sufficiently limit potential bias.**			**Low**
5. Study Confounding		**Goal: To judge the risk of bias due to confounding (where the effect of the PF is distorted by another factor that is related to both the PF and outcome).**			
a)	Important Confounders Measured	Important confounders including treatments are measured (key LIST variables)	None assessed	No	
b)	Definition of the confounding factor	Clear definitions of the important confounders measured are provided (including dose, level, and duration of exposure).	None assessed	No	
c)	Valid and Reliable Measurement of Confounders	Measurement of all important confounders is adequately valid and reliable (may include relevant outside sources of information on measurement	None assessed	No	

		properties, also characteristics, such as blind measurement and limited reliance on recall).		
d)	Method and Setting of Confounding Measurement	The method and setting of confounding measurement are the same for all study participants.	None assessed	No
e)	Method used for missing data	Appropriate methods are used if imputation is used for missing confounder data.	None assessed	No
f)	Appropriate Accounting for Confounding	Important potential confounders are accounted for in the study design (for example, matching for key variables, stratification, or initial assembly of comparable groups).	No matching apparent or stratification	No

	Important potential confounders are accounted for in the analysis (that is, appropriate adjustment).	None assessed	No	
Study Confounding Summary	**Important potential confounders are appropriately accounted for, limiting potential bias with respect to the relationship between PF and outcome.**			**High**
6. Statistical Analysis and Reporting	**Goal: To judge the risk of bias related to the statistical analysis and presentation of results.**			
a) Presentation of analytical strategy	There is sufficient presentation of data to assess the adequacy of the analysis.	Only briefly states t-test value has been used and sets p value for significance at 0.05.	Partial	

b)	Model development strategy	The strategy for model building (inclusion of variables in the statistical model) is appropriate and based on a conceptual framework or model. The selected statistical model is adequate for the design of the study.	Does not build a model and gives no further information on the statistical analysis	No	
c)	Reporting of results	There is no selective reporting of results.	None apparent but unclear	Unsure	
Statistical Analysis and Presentation Summary		**The statistical analysis is appropriate for the design of the study, limiting potential for presentation of invalid or spurious results.**			**High: small groups for comparison and limited data on analysis**
Summary Badet et al 2009: Participation high; Attrition moderate; PF high; Outcome low; Confounding high; Statistical Analysis high					

Table 30.7

Author and year of publication	Koch et al 2001			
Biases	**Issues to consider for judging overall rating of "Risk of bias"**	**Study Methods & Comments**	**Rating of reporting: yes, partial, no, unsure**	**Overall rating of "Risk of bias" for domain: high, moderate, low**
1. Study Participation	**Goal: To judge the risk of selection bias: the likelihood that relationship between prognostic factor (PF) and outcome is different for participants and eligible non-participants. For registry studies, this section has considered the basis for selection for the registry (as all people with CF in the region would theoretically be eligible for analysis) as the equivalent of "participant selection" in a prospective study**			
a) Source of target population	The source population or population of interest is adequately described for key characteristics (LIST).	States that 11,749 were in registry – representing 50% with CF in European countries – 8,963 genotyped (76%). Only gives assessed variables for those genotyped by class which isn't complete for all list variables.	No	
b) Method used to identify	The sampling frame and recruitment are adequately	Explains registry began enrolling 1994 and has data from 9 listed European countries. As above covers 50% with CF. Specific	Partial	

	population	described, including methods to identify the sample sufficient to limit potential bias (for example, referral patterns in health care)	process by which patients are entered into the registry is unclear so unclear whether certain geographic regions may have limited coverage so account for missing data on half of all people with CF.	
c)	Recruitment period	Period of recruitment is adequately described	States enrolment started 1994 with data taken from enrolment forms and ideally up to 4 annual follow-up assessments. Assessment 1997.	Yes
d)	Place of recruitment	Place of recruitment (setting and geographic location) are adequately described	Lists countries and says this covers 50% of those with CF across these countries. Doesn't describe what locations within these countries are covered or whether some countries may give greater representation than others.	Partial
e)	Inclusion and exclusion criteria	Inclusion and exclusion criteria are adequately described (including explicit diagnostic criteria)	Study includes participants genotyped and classified. The registry covers 50% and as above unclear whether there may be less clinic coverage in certain geographic regions accounting for those not entered into the registry	Partial
f)	Adequate	There is adequate	The registry covers only 50% of all people	No

	study participation	participation in the study by eligible individuals (>70%)	with CF. 76% of this cohort are genotyped but the high genotyping rate could be associated with why these people are in the registry.		
g)	Baseline characteristics	The baseline study sample (individuals entering the study) is adequately described for key characteristics (LIST).	No information given for the full sample or for the genotyped sample aside from the assessed phenotypic variables.	No	
Summary Study participation		**The study sample represents the population of interest on key characteristics, sufficient to limit potential bias for the observed relationship between the PF and outcome.**			**High: this registry only covers half with CF in these countries for unclear reasons**
2. Study Attrition		**Goal: To judge the risk of attrition bias (likelihood that relationship between PF and the outcome are different for completing and non-completing participants). For registry studies, we considered that this section should consider loss of participants from the analysis due to lack of available data, for example.**			
a)	Proportion of baseline sample available for	Response rate is adequate (proportion of study sample completing the study	There was no apparent loss of participants among those with genotyping/classification data available. They account for 76% of those in the cohort, though groupings for	Partial	

	analysis	and providing outcome data).	certain class combinations (3/3 and 5/any) are small.		
b)	Attempts to collect information on participants who dropped out	Attempts to collect information on participants who dropped out of the study are described.	No information is available for the 24% without data available.	No	
c)	Reasons and potential impact of subjects lost to follow-up	Reasons for loss to follow-up are provided.	As above it's clear that the study has only included those genotyped/classified but there is no further detail on why participants may not have been genotyped.	Partial	
d)	Outcome and prognostic factor information on those lost to follow-up	Participants lost to follow-up are adequately described for key characteristics (LIST) with no important differences from participants.	No information is available for those who were not genotyped.	No	
Study Attrition Summary		**Loss to follow-up (from baseline sample to study population analysed) is not associated with key characteristics (that is, the study data adequately represent the sample) sufficient to limit potential bias to the observed relationship between PF and outcome.**			**Moderate: higher proportion**

				genotyped that other studies but unclear whether there may be differences for those not genotyped
3. Prognostic Factor Measurement	**Goal: To judge the risk of measurement bias related to how the PF was measured (differential measurement of PF related to the level of outcome). For studies comparing variant classes this includes whether the system used to classify was adequately described.**			
a) Definition of the PF	A clear definition or description of 'PF' is provided (including dose, duration of exposure, and clear specification of the measurement method).	Clearly describes studied variants and classification 1-5.	Yes	
b) Valid and Reliable Measurement of PF	Method of PF measurement is valid and reliable to limit misclassification bias (may include relevant outside	Technical method of genotyping is not given. G85E has since been reclassified. There are some variants that have been added based on similar localisation within the gene to others, which could introduce error.	Partial	

		sources of information on measurement properties, such as blind measurement and limited reliance on recall).			
c)	Method and Setting of PF Measurement	The method and setting of measurement of PF is the same for all study participants.	Genotyping may have differed across different countries and facilities.	Unsure	
d)	Proportion of data on PF available for analysis	Adequate proportion (>70%) of the study sample has complete data for PF variable.	76% of the available cohort genotyped or classified.	Yes	
e)	Method used for missing data	Appropriate methods of imputation are used for missing PF data.	Unclear if any imputation used for genotype data recorded in the registry.	Unsure	
PF Measurement Summary		**PF is adequately measured in study participants to sufficiently limit potential bias.**			**Moderate: mainly due to potential variation in lab methods and possible**

				misclassification
4. Outcome Measurement		**Goal: To judge the risk of bias related to the measurement of outcome (differential measurement of outcome related to the baseline level of PF).**		
a)	Definition of the Outcome	A clear definition of outcome is provided, including duration of follow-up.	Describes observation period for phenotypic assessments is from enrolment to the subsequent 180 days and lists the variables and age groups assessed. Says that first valid input has been used for lung, function, age, BMI. Less clear for other variables.	Partial
b)	Valid and Reliable Measurement of Outcome	The method of outcome measurement used is adequately valid and reliable to limit misclassification bias (may include relevant outside sources of information on measurement properties, also characteristics, such as blind measurement and	FEV1, age and BMI may be less likely to introduce error. Pancreatic sufficiency by ERT as standard. *P. aeruginosa* unclear.	Partial

		confirmation of outcome with valid and reliable test).			
c)	Method and Setting of Outcome Measurement	The method and setting of outcome measurement is the same for all study participants.	This is expected to have varied across centres across countries.	No	
Outcome Measurement Summary		**Outcome of interest is adequately measured in study participants to sufficiently limit potential bias.**			**High: information not available on all variables and as this is across countries may be more discrepancy**
5. Study Confounding		**Goal: To judge the risk of bias due to confounding (where the effect of the PF is distorted by another factor that is related to both the PF and outcome).**			
a)	Important Confounders Measured	Important confounders including treatments are measured (key LIST variables)	Confounders not assessed	No	
b)	Definition of the	Clear definitions of the important	Confounders not assessed	No	

	confounding factor	confounders measured are provided (including dose, level, and duration of exposure).		
c)	Valid and Reliable Measurement of Confounders	Measurement of all important confounders is adequately valid and reliable (may include relevant outside sources of information on measurement properties, also characteristics, such as blind measurement and limited reliance on recall).	Confounders not assessed	No
d)	Method and Setting of Confounding Measureme	The method and setting of confounding measurement are the same for all	Confounders not assessed	No

	nt	study participants.			
e)	Method used for missing data	Appropriate methods are used if imputation is used for missing confounder data.	Not applicable as not measured	No	
f)	Appropriate Accounting for Confoundin g	Important potential confounders are accounted for in the study design (for example, matching for key variables, stratification, or initial assembly of comparable groups).	No stratification.	No	
		Important potential confounders are accounted for in the analysis (that is, appropriate adjustment).	Confounders not assessed	No	
Study Confounding Summary		**Important potential confounders are appropriately accounted for, limiting potential bias with respect to the relationship between PF and outcome.**			**High**
6. Statistical		**Goal: To judge the risk of bias related to the statistical analysis and presentation of results.**			

Analysis and Reporting					
a)	Presentation of analytical strategy	There is sufficient presentation of data to assess the adequacy of the analysis.	No statistical analysis	No	
b)	Model development strategy	The strategy for model building (inclusion of variables in the statistical model) is appropriate and based on a conceptual framework or model. The selected statistical model is adequate for the design of the study.	No statistical analysis	No	
c)	Reporting of results	There is no selective reporting of results.	No statistical analysis	No	
Statistical Analysis and Presentation Summary		**The statistical analysis is appropriate for the design of the study, limiting potential for presentation of invalid or spurious results.**			**NA – statistical analysis not performed, comparison of**

	mean ranges only

Summary Koch et al 2001: Participation high; Attrition moderate; PF moderate; Outcome high; Confounding high; Statistical Analysis NA

Table 30.8

Author and year of publication	**Dewulf et al 2015**			
Biases	**Issues to consider for judging overall rating of "Risk of bias"**	**Study Methods & Comments**	**Rating of reporting: yes, partial, no, unsure**	**Overall rating of "Risk of bias" for domain: high, moderate, low**
1. Study Participation	**Goal: To judge the risk of selection bias: the likelihood that relationship between prognostic factor (PF) and outcome is different for participants and eligible non-participants. For registry studies, this section has considered the basis for selection for the registry (as all people with CF in the region would theoretically be eligible for analysis) as the equivalent of "participant selection" in a prospective study**			
a) Source of target population	The source population or population of interest is adequately described for key	States that 1138 were in registry – representing >90% with CF in Belgium. 75% were genotyped. Only gives assessed variables for those	No	

		characteristics (LIST).	genotyped or classified, not full registry.	
b)	Method used to identify population	The sampling frame and recruitment are adequately described, including methods to identify the sample sufficient to limit potential bias (for example, referral patterns in health care)	The study looks at 2010 and covers genotyped and non-transplanted patients. Registry covers >90% of those with CF in Belgium. Specific process by which patients are entered into the registry is unclear but the high coverage gives greater confidence in representation	Yes
c)	Recruitment period	Period of recruitment is adequately described	Study looks at 2010 though enrolment period for the registry is unclear.	Partial
d)	Place of recruitment	Place of recruitment (setting and geographic location) are adequately described	Doesn't say specifically but Belgium and high coverage so expected to cover most of the country.	Yes
e)	Inclusion and exclusion criteria	Inclusion and exclusion criteria are adequately described (including explicit diagnostic criteria)	Study includes participants genotyped and classified and who haven't received a transplant. The registry covers >90% so unexpected to be exclusions.	Yes
f)	Adequate study	There is adequate participation in the	>90% with CF and study covers 75% of them.	Yes

	participation	study by eligible individuals (>70%)			
g)	Baseline characteristics	The baseline study sample (individuals entering the study) is adequately described for key characteristics (LIST).	Aside from ethnicity, information is given for the full study sample with and without transplant.	Partial	
Summary Study participation		**The study sample represents the population of interest on key characteristics, sufficient to limit potential bias for the observed relationship between the PF and outcome.**			**Moderate: on basis doesn't give characteristics for the full registry**
2. Study Attrition		**Goal: To judge the risk of attrition bias (likelihood that relationship between PF and the outcome are different for completing and non-completing participants). For registry studies, we considered that this section should consider loss of participants from the analysis due to lack of available data, for example.**			
a)	Proportion of baseline sample available for analysis	Response rate is adequate (proportion of study sample completing the study and providing outcome data).	Those with genotyping classification data account for 75% of cohort, further exclusion of transplant patients takes participation to 66%	Partial	
b)	Attempts to	Attempts to collect	Characteristics (comparing mild/severe	Partial	

	collect information on participants who dropped out	information on participants who dropped out of the study are described.	groups) have been given for all 853 genotyped including transplant patients, then for the 748 without transplant. No information is available for the 25% without genotyping/classification.		
c)	Reasons and potential impact of subjects lost to follow-up	Reasons for loss to follow-up are provided.	The study has only included those genotyped and classified and those without transplant.	Yes	
d)	Outcome and prognostic factor information on those lost to follow-up	Participants lost to follow-up are adequately described for key characteristics (LIST) with no important differences from participants.	As above characteristics for those including and excluding transplant (though without direct comparison) but nothing for the full registry cohort.	Partial	
Study Attrition Summary		**Loss to follow-up (from baseline sample to study population analysed) is not associated with key characteristics (that is, the study data adequately represent the sample) sufficient to limit potential bias to the observed relationship between PF and outcome.**			**Moderate: primarily due to lack of comparison to non-genotyped**

				cohort
3. Prognostic Factor Measurement		**Goal: To judge the risk of measurement bias related to how the PF was measured (differential measurement of PF related to the level of outcome). For studies comparing variant classes this includes whether the system used to classify was adequately described.**		
a)	Definition of the PF	A clear definition or description of 'PF' is provided (including dose, duration of exposure, and clear specification of the measurement method).	Clearly describes studied variants and classification 1-5 using established system.	Yes
b)	Valid and Reliable Measurement of PF	Method of PF measurement is valid and reliable to limit misclassification bias (may include relevant outside sources of information on measurement properties, such as blind measurement and limited reliance on recall).	The study is recent and uses the most up-to-date system.	Yes
c)	Method and	The method and	Genotyping may have differed across	Unsure

	Setting of PF Measurement	setting of measurement of PF is the same for all study participants.	facilities in Belgium but this is expected to be minimal as this is one country.		
d)	Proportion of data on PF available for analysis	Adequate proportion (>70%) of the study sample has complete data for PF variable.	75% of the available cohort genotyped or classified. Further exclusion due to transplant reduced the proportion but that is appropriate exclusion for purpose of analysis.	Yes	
e)	Method used for missing data	Appropriate methods of imputation are used for missing PF data.	Unclear if any imputation used for genotype data recorded in the registry.	Unsure	
PF Measurement Summary		**PF is adequately measured in study participants to sufficiently limit potential bias.**			**Low**
4. Outcome Measurement		**Goal: To judge the risk of bias related to the measurement of outcome (differential measurement of outcome related to the baseline level of PF).**			
a)	Definition of the Outcome	A clear definition of outcome is provided, including duration of follow-up.	Lists the treatments and assesses use during a one-year period. Clearly lists how other variables were assessed.	Yes	
b)	Valid and Reliable Measurement of	The method of outcome measurement used is adequately valid and	Main outcome of treatment burden index is only an estimate and based on medications listed in patient charts. This is likely the best objective estimate of treatment burden	Unsure	

	Outcome	reliable to limit misclassification bias (may include relevant outside sources of information on measurement properties, also characteristics, such as blind measurement and confirmation of outcome with valid and reliable test).	though uncertain whether there could be inaccuracies Pancreatic sufficiency assessed by stool fat content and fecal elastase. *P. aeruginosa* colonisation by defined criteria.		
c)	Method and Setting of Outcome Measurement	The method and setting of outcome measurement is the same for all study participants.	This is expected to have varied across centres within the country but may be less variation than other multicentre studies.	Unsure	
Outcome Measurement Summary		**Outcome of interest is adequately measured in study participants to sufficiently limit potential bias.**			**Moderate: treatment burden is likely to be the best objective estimate available but unclear whether there**

					could be error
5. Study Confounding		**Goal: To judge the risk of bias due to confounding (where the effect of the PF is distorted by another factor that is related to both the PF and outcome).**			
a)	Important Confounders Measured	Important confounders including treatments are measured (key LIST variables)	Age, gender and FEV1 for treatment analysis. Ethnicity is missing. Study is recent and is assessing treatment. Method of diagnosis isn't adjusted though this is non-screen setting. Doesn't adjust for confounders in other analyses.	Partial No	For treatment analysis only Other phenotypic variables
b)	Definition of the confounding factor	Clear definitions of the important confounders measured are provided (including dose, level, and duration of exposure).	Age, gender, FEV1	Partial No	For treatment analysis only Other phenotypic variables
c)	Valid and Reliable Measurement of Confounders	Measurement of all important confounders is adequately valid and reliable (may include relevant outside sources of information on measurement	Age, gender, FEV1 likely to be variable	Partial No	For treatment analysis only Other phenotypic variables

		properties, also characteristics, such as blind measurement and limited reliance on recall).			
d)	Method and Setting of Confounding Measurement	The method and setting of confounding measurement are the same for all study participants.	Not expected to be error in assessed variables	Partial No	For treatment analysis only Other phenotypic variables
e)	Method used for missing data	Appropriate methods are used if imputation is used for missing confounder data.	Unclear	Unsure	
f)	Appropriate Accounting for Confounding	Important potential confounders are accounted for in the study design (for example, matching for key variables, stratification, or initial assembly of comparable groups).	No stratification.	No	

	Important potential confounders are accounted for in the analysis (that is, appropriate adjustment).	Confounders not assessed	Partial No	For treatment analysis only Other phenotypic variables
Study Confounding Summary	**Important potential confounders are appropriately accounted for, limiting potential bias with respect to the relationship between PF and outcome.**			**Moderate: treatment High: other variables**
6. Statistical Analysis and Reporting	**Goal: To judge the risk of bias related to the statistical analysis and presentation of results.**			
a) Presentation of analytical strategy	There is sufficient presentation of data to assess the adequacy of the analysis.	Multiple regression model for treatment burden. Chi-squared or Cochrane-Mantel-Haenszel for categorical data (Fisher's for small numbers), Mann-Whitney for continuous.	Yes	

b)	Model development strategy	The strategy for model building (inclusion of variables in the statistical model) is appropriate and based on a conceptual framework or model. The selected statistical model is adequate for the design of the study.	Multiple regression to account for confounding. States p<0.05 considered significant.	Yes	
c)	Reporting of results	There is no selective reporting of results.	None apparent.	No	
Statistical Analysis and Presentation Summary		**The statistical analysis is appropriate for the design of the study, limiting potential for presentation of invalid or spurious results.**			**Low**

Summary Dewulf et al 2015: Participation moderate; Attrition moderate; PF low; Outcome moderate; Confounding moderate treatment/high other; Statistical Analysis low

Table 30.9

Author and year of publication	Green et al 2010			
Biases	**Issues to consider for judging overall rating of "Risk of bias"**	**Study Methods & Comments**	**Rating of reporting: yes, partial, no, unsure**	**Overall rating of "Risk of bias" for domain: high, moderate, low**
1. Study Participation	**Goal: To judge the risk of selection bias: the likelihood that relationship between prognostic factor (PF) and outcome is different for participants and eligible non-participants. For registry studies, this section has considered the basis for selection for the registry (as all people with CF in the region would theoretically be eligible for analysis) as the equivalent of "participant selection" in a prospective study**			
a) Source of target population	The source population or population of interest is adequately described for key characteristics (LIST).	US Twin and Sibling Study recruited on basis of having a surviving twin/sibling with CF. Assessed participants represent 83% of this cohort. Most characteristics of these participants given, though not for full potential sample.	Partial	
b) Method used to identify	The sampling frame and recruitment are adequately described,	As above twin/sibling and says 99% attending centres in the US. Further details on recruitment or how representative this	Partial	

	population	including methods to identify the sample sufficient to limit potential bias (for example, referral patterns in health care)	sample is are unclear.	
c)	Recruitment period	Period of recruitment is adequately described	This study said to include those with at least annual microbiology assessments up to 2008 but baseline period and study recruitment is unclear from this study.	No
d)	Place of recruitment	Place of recruitment (setting and geographic location) are adequately described	US but isn't clear how representative the twin/sibling study is.	No
e)	Inclusion and exclusion criteria	Inclusion and exclusion criteria are adequately described (including explicit diagnostic criteria)	Study includes participants genotyped and classified and with infection data. Unclear how representative the twin/sibling study is.	Partial
f)	Adequate study participation	There is adequate participation in the study by eligible individuals (>70%)	The analysis represents 83% of the cohort but unclear how representative they are of all potentially eligible for the twin/sibling study.	Unsure
g)	Baseline characteristi	The baseline study sample (individuals	Age, gender, ethnicity, baseline characteristics are given for those with	Partial

	cs	entering the study) is adequately described for key characteristics (LIST).	mild/severe genotypes.		
Summary Study participation		**The study sample represents the population of interest on key characteristics, sufficient to limit potential bias for the observed relationship between the PF and outcome.**			**High: on basis that the study is representative of those with sibling with CF only**
2. Study Attrition		**Goal: To judge the risk of attrition bias (likelihood that relationship between PF and the outcome are different for completing and non-completing participants). For registry studies, we considered that this section should consider loss of participants from the analysis due to lack of available data, for example.**			
a)	Proportion of baseline sample available for analysis	Response rate is adequate (proportion of study sample completing the study and providing outcome data).	Those with genotyping/classification and infection data account for 83% of the cohort	Yes	
b)	Attempts to collect information on	Attempts to collect information on participants who dropped out of the	Characteristics are not given for those not included in the analysis.	No	

	participants who dropped out	study are described.			
c)	Reasons and potential impact of subjects lost to follow-up	Reasons for loss to follow-up are provided.	The study has only included those genotyped and classified and with sufficient infection data.	Yes	
d)	Outcome and prognostic factor information on those lost to follow-up	Participants lost to follow-up are adequately described for key characteristics (LIST) with no important differences from participants.	Characteristics are only listed for those included in the analysis.	No	
Study Attrition Summary		**Loss to follow-up (from baseline sample to study population analysed) is not associated with key characteristics (that is, the study data adequately represent the sample) sufficient to limit potential bias to the observed relationship between PF and outcome.**			**Moderate: primarily due to lack of comparison to non-genotyped cohort**
3. Prognostic		**Goal: To judge the risk of measurement bias related to how the PF was measured (differential**			

Factor		measurement of PF related to the level of outcome). For studies comparing variant classes this includes whether the system used to classify was adequately described.		
Measurement				
a)	Definition of the PF	A clear definition or description of 'PF' is provided (including dose, duration of exposure, and clear specification of the measurement method).	Lists the variants that have been analysed in groups 1 to 5.	Yes
b)	Valid and Reliable Measurement of PF	Method of PF measurement is valid and reliable to limit misclassification bias (may include relevant outside sources of information on measurement properties, such as blind measurement and limited reliance on recall).	Doesn't explicitly describe what system has been used. References McKone 2003 and other earlier publications though there are differences and additions from the De Boeck 14 and Mckone 06 listings.	Unsure
c)	Method and Setting of PF Measureme	The method and setting of measurement of PF is the same for all study	Genotyping may have differed across facilities in the US.	Unsure

	nt	participants.			
d)	Proportion of data on PF available for analysis	Adequate proportion (>70%) of the study sample has complete data for PF variable.	83% of the available cohort genotyped or classified. Further exclusion due to transplant reduced the proportion but that is appropriate exclusion for purpose of analysis.	Yes	
e)	Method used for missing data	Appropriate methods of imputation are used for missing PF data.	Unclear if any imputation used for genotype data recorded in the registry.	Unsure	
PF Measurement Summary		**PF is adequately measured in study participants to sufficiently limit potential bias.**			**Moderate: genotyping for the majority but uncertainties around classification system**
4. Outcome Measurement		**Goal: To judge the risk of bias related to the measurement of outcome (differential measurement of outcome related to the baseline level of PF).**			
a)	Definition of the Outcome	A clear definition of outcome is provided, including duration of follow-up.	Clear definition for infection outcomes Other variables FEV1 unlikely to be biased but unclear how pancreatic status is assessed	Yes Partial	Infection Other variables
b)	Valid and Reliable	The method of outcome measurement	Uses four different definitions of infection status.	Yes Partial	Infection Other

	Measurement of Outcome	used is adequately valid and reliable to limit misclassification bias (may include relevant outside sources of information on measurement properties, also characteristics, such as blind measurement and confirmation of outcome with valid and reliable test).	Other phenotypic variables are taken primarily as baseline characteristics and limited information on assessment is given.		variables.
c)	Method and Setting of Outcome Measurement	The method and setting of outcome measurement is the same for all study participants.	This is expected to have varied across centres but definitions are clearly given for infection variables which should limit misclassification. Uncertainty around other variables	Yes Unsure.	Infection Other variables.
Outcome Measurement Summary		**Outcome of interest is adequately measured in study participants to sufficiently limit potential bias.**			**Low: infection High: other variables due to limited information**
5. Study Confounding		**Goal: To judge the risk of bias due to confounding (where the effect of the PF is distorted by another factor that is related to both the PF and outcome).**			

a)	Important Confounders Measured	Important confounders including treatments are measured (key LIST variables)	Infection is appropriately adjusted for FEV1 in the year prior to analysis and number of cultures. Gender, ethnicity and pancreatic status had also initially been assessed in univariate regression analysis but as they didn't have significant effect weren't included in the final model. Treatment/year of entry and age/method of diagnosis not assessed. Other assessments are baseline characteristics with no information	Partial No	Infection Other variables
b)	Definition of the confounding factor	Clear definitions of the important confounders measured are provided (including dose, level, and duration of exposure).	Yes for FEV1 in the year prior to analysis and number of cultures. Other variables no adjustment.	Yes No	Treatment Other
c)	Valid and Reliable Measurement of Confounders	Measurement of all important confounders is adequately valid and reliable (may include relevant outside sources of information on measurement	For those assessed.	Yes No	Treatment Other

		properties, also characteristics, such as blind measurement and limited reliance on recall).		
d)	Method and Setting of Confounding Measurement	The method and setting of confounding measurement are the same for all study participants.	Those assessed are likely in different settings but no other variables assessed.	Unsure
e)	Method used for missing data	Appropriate methods are used if imputation is used for missing confounder data.	Unclear	Unsure
f)	Appropriate Accounting for Confounding	Important potential confounders are accounted for in the study design (for example, matching for key variables, stratification, or initial assembly of comparable groups).	No stratification.	No

	Important potential confounders are accounted for in the analysis (that is, appropriate adjustment).	For those assessed.	Yes	
Study Confounding Summary	**Important potential confounders are appropriately accounted for, limiting potential bias with respect to the relationship between PF and outcome.**			**Moderate: infection High: other variables**
6. Statistical Analysis and Reporting	**Goal: To judge the risk of bias related to the statistical analysis and presentation of results.**			
a) Presentation of analytical strategy	There is sufficient presentation of data to assess the adequacy of the analysis.	Multivariate Cox regression model for infection	Yes	

b)	Model development strategy	The strategy for model building (inclusion of variables in the statistical model) is appropriate and based on a conceptual framework or model. The selected statistical model is adequate for the design of the study.	Explanation of univariate and multivariate Cox regression model is given alongside p values for significance	Yes	
c)	Reporting of results	There is no selective reporting of results.	None apparent.	No	
Statistical Analysis and Presentation Summary		**The statistical analysis is appropriate for the design of the study, limiting potential for presentation of invalid or spurious results.**			**Low**

Summary Green et al 2010: Participation high; Attrition moderate; PF moderate; Outcome low infection/high others; Confounding moderate infection/high others; Statistical Analysis low

Table 30.10

Author and year of publication	Radtke et al 2017			
Biases	**Issues to consider for judging overall rating of "Risk of bias"**	**Study Methods & Comments**	**Rating of reporting: yes, partial, no, unsure**	**Overall rating of "Risk of bias" for domain: high, moderate, low**
1. Study Participation	**Goal: To judge the risk of selection bias: the likelihood that relationship between prognostic factor (PF) and outcome is different for participants and eligible non-participants. For registry studies, this section has considered the basis for selection for the registry (as all people with CF in the region would theoretically be eligible for analysis) as the equivalent of "participant selection" in a prospective study**			
a) Source of target population	The source population or population of interest is adequately described for key characteristics (LIST).	Lists 17 countries covered representative of 32 asked to give data on ≥20 patients, aged ≥8 years who completed a maximal cardiopulmonary exercise test (CPET), No data on the full eligible population though and unclear how representative these 17 countries are.	No	
b) Method used to identify	The sampling frame and recruitment are adequately described,	Describes criteria as above but unclear who may have been eligible for CPET within centres or whether the centre's selection of	Partial	

	population	including methods to identify the sample sufficient to limit potential bias (for example, referral patterns in health care)	>20 participants was representative.	
c)	Recruitment period	Period of recruitment is adequately described	Assessments completed1999 to 2014	Yes
d)	Place of recruitment	Place of recruitment (setting and geographic location) are adequately described	Included countries and numbers of centres are listed. Also gives reason for non-participation of the remainder. But unclear how geographically representative the included centres are.	Partial
e)	Inclusion and exclusion criteria	Inclusion and exclusion criteria are adequately described (including explicit diagnostic criteria)	As above for participation. The study further excluded from the analysis people with missing genotype data, exercise data, those aged <12 and who didn't reach their maximal exercise capacity.	Yes
f)	Adequate study participation	There is adequate participation in the study by eligible individuals (>70%)	The centres represent only 50% of those asked, and unclear how representative their patients are.	No
g)	Baseline characteristi	The baseline study sample (individuals	Age, gender, baseline characteristics are given for those with valid exercise data and	Partial

	cs	entering the study) is adequately described for key characteristics (LIST).	not. Country, genotype or ethnicity not compared, nor data given for the full study cohort.		
Summary Study participation		**The study sample represents the population of interest on key characteristics, sufficient to limit potential bias for the observed relationship between the PF and outcome.**			**High: on basis that study has uncertain representation of all with CF**
2. Study Attrition		**Goal: To judge the risk of attrition bias (likelihood that relationship between PF and the outcome are different for completing and non-completing participants). For registry studies, we considered that this section should consider loss of participants from the analysis due to lack of available data, for example.**			
a)	Proportion of baseline sample available for analysis	Response rate is adequate (proportion of study sample completing the study and providing outcome data).	88% with genotyping data but further exclusions due to lack of CPET or other data giving final representation of 73%	Partial	
b)	Attempts to collect information on participants who	Attempts to collect information on participants who dropped out of the study are described.	Characteristics are given comparing those with maximal CPET data to the n=112 without but no information for n=152 excluded for lack of genotyping or other data	Partial	

	dropped out				
c)	Reasons and potential impact of subjects lost to follow-up	Reasons for loss to follow-up are provided.	The study lists reasons for exclusion	Yes	
d)	Outcome and prognostic factor information on those lost to follow-up	Participants lost to follow-up are adequately described for key characteristics (LIST) with no important differences from participants.	Characteristics are compared for those with maximal CPET data and not. Those with missing data were older with higher infection rates and lower FEV1. Unclear whether there may have been differences in genotype. No comparison to others not included.	No	
Study Attrition Summary		**Loss to follow-up (from baseline sample to study population analysed) is not associated with key characteristics (that is, the study data adequately represent the sample) sufficient to limit potential bias to the observed relationship between PF and outcome.**			**High: on basis of differences in those reaching maximal exercise capacity or not**
3. Prognostic Factor		**Goal: To judge the risk of measurement bias related to how the PF was measured (differential measurement of PF related to the level of outcome). For studies comparing variant classes this**			

Measurement		**includes whether the system used to classify was adequately described.**		
a)	Definition of the PF	A clear definition or description of 'PF' is provided (including dose, duration of exposure, and clear specification of the measurement method).	Lists the variants that have been analysed in groups 1 to 5.	Yes
b)	Valid and Reliable Measurement of PF	Method of PF measurement is valid and reliable to limit misclassification bias (may include relevant outside sources of information on measurement properties, such as blind measurement and limited reliance on recall).	Doesn't explicitly describe what classification list has been used. References McKone 2006 but differences and additions.	Unsure
c)	Method and Setting of PF Measurement	The method and setting of measurement of PF is the same for all study participants.	Genotyping may have differed across countries. Classification was by a geneticist blinded to exercise data.	Partial

d)	Proportion of data on PF available for analysis	Adequate proportion (>70%) of the study sample has complete data for PF variable.	88% of the available cohort genotyped or classified. Further exclusion were for other reasons as above	Yes	
e)	Method used for missing data	Appropriate methods of imputation are used for missing PF data.	Unclear if any imputation used for genotype data recorded in the registry.	Unsure	
PF Measurement Summary		**PF is adequately measured in study participants to sufficiently limit potential bias.**			**Moderate: potential variation in genotyping across countries and some uncertainties around classification**
4. Outcome Measurement		**Goal: To judge the risk of bias related to the measurement of outcome (differential measurement of outcome related to the baseline level of PF).**			
a)	Definition of the Outcome	A clear definition of outcome is provided, including duration of follow-up.	Most variables described though some unclear, such as pancreatic status. These were also only baseline characteristics rather than the main aim of the study which was to look at exercise capacity.	Partial	
b)	Valid and	The method of	Age, BMI, FEV1, *P. aeruginosa* described.	Partial	

	Reliable Measurement of Outcome	outcome measurement used is adequately valid and reliable to limit misclassification bias (may include relevant outside sources of information on measurement properties, also characteristics, such as blind measurement and confirmation of outcome with valid and reliable test).	Less clear how pancreatic status was assessed.		
c)	Method and Setting of Outcome Measurement	The method and setting of outcome measurement is the same for all study participants.	Likely to be variability across centres.	No	
Outcome Measurement Summary		**Outcome of interest is adequately measured in study participants to sufficiently limit potential bias.**			**Moderate: these were baseline characteristics but no clear indication of**

					inaccuracy (excluding PS/I measure)
5. Study Confounding		**Goal: To judge the risk of bias due to confounding (where the effect of the PF is distorted by another factor that is related to both the PF and outcome).**			
a)	Important Confounders Measured	Important confounders including treatments are measured (key LIST variables)	Assessed for exercise capacity only, not other variables	No	
b)	Definition of the confounding factor	Clear definitions of the important confounders measured are provided (including dose, level, and duration of exposure).	Not assessed	No	
c)	Valid and Reliable Measurement of Confounders	Measurement of all important confounders is adequately valid and reliable (may include relevant outside sources of information on measurement properties, also characteristics, such	Not assessed	No	

		as blind measurement and limited reliance on recall).		
d)	Method and Setting of Confounding Measurement	The method and setting of confounding measurement are the same for all study participants.	Not assessed	No
e)	Method used for missing data	Appropriate methods are used if imputation is used for missing confounder data.	Not assessed	No
f)	Appropriate Accounting for Confounding	Important potential confounders are accounted for in the study design (for example, matching for key variables, stratification, or initial assembly of comparable groups).	No stratification.	No
		Important potential confounders are accounted for in the analysis (that is,	Not assessed	No

	appropriate adjustment).			
Study Confounding Summary	**Important potential confounders are appropriately accounted for, limiting potential bias with respect to the relationship between PF and outcome.**			**High**
6. Statistical Analysis and Reporting	**Goal: To judge the risk of bias related to the statistical analysis and presentation of results.**			
a) Presentation of analytical strategy	There is sufficient presentation of data to assess the adequacy of the analysis.	Analysis of variance, Chi squared, Kruskal-Wallis to compare variables between groups. Multivariate model only assessed for exercise capacity	Partial	
b) Model development strategy	The strategy for model building (inclusion of variables in the statistical model) is appropriate and based on a conceptual framework or model. The selected statistical model is adequate for the design of the	As above, primary aim was to analyse factors associated with exercise capacity rather than other phenotypic variables. Tests comparing variables seem appropriate but p values not given	Partial	

		study.			
c)	Reporting of results	There is no selective reporting of results.	None apparent.	No	
Statistical Analysis and Presentation Summary		**The statistical analysis is appropriate for the design of the study, limiting potential for presentation of invalid or spurious results.**			**Moderate: primary design of model was not to assess other variables**
Summary Radtke et al 2017: Participation high; Attrition high; PF moderate; Outcome moderate; Confounding high; Statistical Analysis moderate					

Table 30.11

Author and year of publication	The Cystic Fibrosis Genotype-Phenotype Consortium 1993			
Biases	**Issues to consider for judging overall rating of "Risk of bias"**	**Study Methods & Comments**	**Rating of reporting: yes, partial, no, unsure**	**Overall rating of "Risk of bias" for domain: high, moderate, low**
1. Study Participation	**Goal: To judge the risk of selection bias: the likelihood that relationship between prognostic factor (PF) and outcome is different for participants and eligible non-participants. For registry studies, this section has considered the basis for selection for the registry (as all people with CF in the region would theoretically be eligible for analysis) as the equivalent of "participant selection" in a prospective study**			
a) Source of target population	The source population or population of interest is adequately described for key characteristics (LIST).	32 centres participated of 89 belonging to the CF Genetic Analysis Consortium. Unclear on the remaining two-thirds of centres so unclear how representative they are	No	
b) Method used to identify	The sampling frame and recruitment are adequately described,	This study invited 89 centres to take part and included age- and sex-matched homozygotes and heterozygotes from the	No	

	population	including methods to identify the sample sufficient to limit potential bias (for example, referral patterns in health care)	same centre (to control for treatment received). Unclear from this paper what coverage the consortium has or how representative these 32 centres are	
c)	Recruitment period	Period of recruitment is adequately described	Unclear assessment period	No
d)	Place of recruitment	Place of recruitment (setting and geographic location) are adequately described	Included centres are listed but unclear which countries/centres are included in the full consortium and what coverage this has	Partial
e)	Inclusion and exclusion criteria	Inclusion and exclusion criteria are adequately described (including explicit diagnostic criteria)	As above for inclusion of age- and sex-matched homozygotes and heterozygotes from the same centre. Says that the consortium includes those with large numbers of genotyped patients but unclear how representative they are.	No
f)	Adequate study participation	There is adequate participation in the study by eligible individuals (>70%)	The centres represent only 36% of those asked, and unclear how representative their patients are.	No
g)	Baseline	The baseline study	Age, gender, baseline characteristics, age at	Yes

characteristics	sample (individuals entering the study) is adequately described for key characteristics (LIST).	diagnosis given for genotypes. Countries of origin listed and states that all were of White ethnicity.		
Summary Study participation	**The study sample represents the population of interest on key characteristics, sufficient to limit potential bias for the observed relationship between the PF and outcome.**			**High: the study includes a third of eligible centres and there are various uncertainties around recruitment period**
2. Study Attrition	**Goal: To judge the risk of attrition bias (likelihood that relationship between PF and the outcome are different for completing and non-completing participants). For registry studies, we considered that this section should consider loss of participants from the analysis due to lack of available data, for example.**			
a) Proportion of baseline sample available for	Response rate is adequate (proportion of study sample completing the study	69% of the included centres genotyped 100% of their patients, 6% genotyped 75% and 25% of centres didn't specify. Otherwise there's a lack of clarity on whether others	Unsure	

	analysis	and providing outcome data).	eligible may not have participated.		
b)	Attempts to collect information on participants who dropped out	Attempts to collect information on participants who dropped out of the study are described.	The assessed phenotypic variables are only given for those genotyped and studied.	No	
c)	Reasons and potential impact of subjects lost to follow-up	Reasons for loss to follow-up are provided.	As above there's limited information and it's unclear whether the centres may have been selective in their patient inclusions.	Unsure	
d)	Outcome and prognostic factor information on those lost to follow-up	Participants lost to follow-up are adequately described for key characteristics (LIST) with no important differences from participants.	As above phenotypic variables are only given for those genotyped and studied.	No	
Study Attrition Summary		**Loss to follow-up (from baseline sample to study population analysed) is not associated with key characteristics (that is, the study data adequately represent the sample) sufficient to limit potential bias to the observed relationship between PF and outcome.**			**High: due to uncertainties around**

					genotyping and inclusion across centres
3. Prognostic Factor Measurement		**Goal: To judge the risk of measurement bias related to how the PF was measured (differential measurement of PF related to the level of outcome). For studies comparing variant classes this includes whether the system used to classify was adequately described.**			
a)	Definition of the PF	A clear definition or description of 'PF' is provided (including dose, duration of exposure, and clear specification of the measurement method).	Lists the variants that have been studied	Yes	
b)	Valid and Reliable Measurement of PF	Method of PF measurement is valid and reliable to limit misclassification bias (may include relevant outside sources of information on measurement properties, such as blind measurement	Doesn't describe genotyping method but lists that the study has paired F508del homozygotes with people with the next 7 most common variants: G542X, R553X, W1282X, N1303K, R117H, 621+1G>T, 1717-1G>A	Partial	

		and limited reliance on recall).			
c)	Method and Setting of PF Measurement	The method and setting of measurement of PF is the same for all study participants.	Genotyping may have differed across countries and facilities.	Unsure	
d)	Proportion of data on PF available for analysis	Adequate proportion (>70%) of the study sample has complete data for PF variable.	Says that all participants across centres were offered genotyping. 69% of the included centres genotyped 100% of their patients and 6% genotyped 75%. Uncertain for the rest but on this basis expected to be >70% coverage.	Yes	
e)	Method used for missing data	Appropriate methods of imputation are used for missing PF data.	Unclear if any imputation used for genotype data recorded in the registry.	Unsure	
PF Measurement Summary		**PF is adequately measured in study participants to sufficiently limit potential bias.**			**Moderate: potential variation in genotyping across countries and some uncertainties around representation**

4. Outcome Measurement		**Goal: To judge the risk of bias related to the measurement of outcome (differential measurement of outcome related to the baseline level of PF).**		
a)	Definition of the Outcome	A clear definition of outcome is provided, including duration of follow-up.	Variables are clearly reported. These appear to be single cross sectional entries for each person though there is some lack of clarity around the assessment period.	Partial
b)	Valid and Reliable Measurement of Outcome	The method of outcome measurement used is adequately valid and reliable to limit misclassification bias (may include relevant outside sources of information on measurement properties, also characteristics, such as blind measurement and confirmation of outcome with valid and reliable test).	Valid descriptions are given on how lung function, *P. aeruginosa* and pancreatic status were assessed.	Yes
c)	Method and Setting of Outcome Measureme	The method and setting of outcome measurement is the same for all study	Likely to be variability across centres – in particular for pancreatic assessment. Centres were asked to report sufficiency or not but used variable methods to assess	No

	nt	participants.	this.	
Outcome Measurement Summary		**Outcome of interest is adequately measured in study participants to sufficiently limit potential bias.**		**Moderate: uncertain assessment period and pancreatic assessment**
5. Study Confounding		**Goal: To judge the risk of bias due to confounding (where the effect of the PF is distorted by another factor that is related to both the PF and outcome).**		
a)	Important Confounders Measured	Important confounders including treatments are measured (key LIST variables)	Age, gender and treatment centre to account for variation in care levels	Partial
b)	Definition of the confounding factor	Clear definitions of the important confounders measured are provided (including dose, level, and duration of exposure).	Limited applicability for age, gender, centre	NA
c)	Valid and Reliable Measurement of Confounders	Measurement of all important confounders is adequately valid and reliable (may include relevant outside sources of information	Limited applicability for age, gender, centre	NA

		on measurement properties, also characteristics, such as blind measurement and limited reliance on recall).		
d)	Method and Setting of Confounding Measurement	The method and setting of confounding measurement are the same for all study participants.	Participants from the same centre	Yes
e)	Method used for missing data	Appropriate methods are used if imputation is used for missing confounder data.	Unclear if any used	Unsure
f)	Appropriate Accounting for Confounding	Important potential confounders are accounted for in the study design (for example, matching for key variables, stratification, or initial assembly of comparable groups).	Matched for age, gender and centre	Yes

		Important potential confounders are accounted for in the analysis (that is, appropriate adjustment).	NA	
Study Confounding Summary		**Important potential confounders are appropriately accounted for, limiting potential bias with respect to the relationship between PF and outcome.**		**Moderate**
6. Statistical Analysis and Reporting		**Goal: To judge the risk of bias related to the statistical analysis and presentation of results.**		
a)	Presentation of analytical strategy	There is sufficient presentation of data to assess the adequacy of the analysis.	Describes two-tail paired t-test for continuous variables and logistic regression for categorical variables.	Yes

b)	Model development strategy	The strategy for model building (inclusion of variables in the statistical model) is appropriate and based on a conceptual framework or model. The selected statistical model is adequate for the design of the study.	The study was not building a statistical model but describes detail as above. There were however small samples. It also reports nominal significance at p=0.05 but then says due to small comparisons “only small probability values interpreted as significant”.	Partial	
c)	Reporting of results	There is no selective reporting of results.	None apparent.	No	
Statistical Analysis and Presentation Summary		**The statistical analysis is appropriate for the design of the study, limiting potential for presentation of invalid or spurious results.**			**Moderate**

Summary CF G-P Consortium 1993: Participation high; Attrition high; PF moderate; Outcome moderate; Confounding moderate; Statistical Analysis moderate

Table 30.12

Author and year of publication	Szczesniak et al et al 2017			
Biases	**Issues to consider for judging overall rating of "Risk of bias"**	**Study Methods & Comments**	**Rating of reporting: yes, partial, no, unsure**	**Overall rating of "Risk of bias" for domain: high, moderate, low**
1. Study Participation	**Goal: To judge the risk of selection bias: the likelihood that relationship between prognostic factor (PF) and outcome is different for participants and eligible non-participants. For registry studies, this section has considered the basis for selection for the registry (as all people with CF in the region would theoretically be eligible for analysis) as the equivalent of "participant selection" in a prospective study**			
a) Source of target population	The source population or population of interest is adequately described for key characteristics (LIST).	Reports characteristics of those included in the study but unclear how representative these people are of the full registry population. Also unclear how representative the registry is of all people with CF in the US.	No	
b) Method used to identify	The sampling frame and recruitment are adequately described,	This study includes patients aged 6-21 years and with FEV1 data collected during the observation period. Only the baseline study	Partial	

	population	including methods to identify the sample sufficient to limit potential bias (for example, referral patterns in health care)	sample is included. Unclear how representative these people are of all in the registry. Also uncertain representation within the registry of all with CF in the US.	
c)	Recruitment period	Period of recruitment is adequately described	Study observation period for the registry is given (1997-2013)	Yes
d)	Place of recruitment	Place of recruitment (setting and geographic location) are adequately described	Registry setting and location is given (US) but unclear which clinics or geographical regions this covers	Partial
e)	Inclusion and exclusion criteria	Inclusion and exclusion criteria are adequately described (including explicit diagnostic criteria)	Study includes those aged 6-21 with FEV1 data collected during the observation period and without transplant. But it's unclear how representative these people are of all in the registry and whether people may have been excluded if lacking genotyping information etc. Also potential for survivor bias.	Unsure
f)	Adequate study participation	There is adequate participation in the study by eligible individuals (>70%)	This study population includes n=18,387 all of whom have been genotyped. Unclear how representative they are of all in the registry or how representative the registry is of all people with CF from this study.	Unsure
g)	Baseline	The baseline study	Age, birth cohort, gender, genotype and	Partial

	characteristics	sample (individuals entering the study) is adequately described for key characteristics (LIST).	baseline characteristics assessed by phenotype (early, middle, late FEV1 decline). Shows gender is associated with lung function decline.		
Summary Study participation		**The study sample represents the population of interest on key characteristics, sufficient to limit potential bias for the observed relationship between the PF and outcome.**			**High: primarily due to lack of clarity on how representative this sample is and potential survivor bias**
2. Study Attrition		**Goal: To judge the risk of attrition bias (likelihood that relationship between PF and the outcome are different for completing and non-completing participants). For registry studies, we considered that this section should consider loss of participants from the analysis due to lack of available data, for example.**			
a)	Proportion of baseline sample available for analysis	Response rate is adequate (proportion of study sample completing the study and providing outcome data).	The population sample reported all have data on the number of F508del copies they are carrying and lung function variables. Unclear how many have been excluded because they didn't have this data.	Unsure	
b)	Attempts to collect	Attempts to collect information on	No coverage of those not included.	Unsure	

	information on participants who dropped out	participants who dropped out of the study are described.			
c)	Reasons and potential impact of subjects lost to follow-up	Reasons for loss to follow-up are provided.	The study states including those with lung function available during the study period. Unclear whether participants may have been excluded due to lack of genotyping or other factors.	Unsure	
d)	Outcome and prognostic factor information on those lost to follow-up	Participants lost to follow-up are adequately described for key characteristics (LIST) with no important differences from participants.	Only characteristics reported for those entering the study.	No	
Study Attrition Summary		**Loss to follow-up (from baseline sample to study population analysed) is not associated with key characteristics (that is, the study data adequately represent the sample) sufficient to limit potential bias to the observed relationship between PF and outcome.**			**High**
3. Prognostic Factor Measurement		**Goal: To judge the risk of measurement bias related to how the PF was measured (differential measurement of PF related to the level of outcome). For studies comparing variant classes this includes whether the system used to classify was adequately described.**			

a)	Definition of the PF	A clear definition or description of 'PF' is provided (including dose, duration of exposure, and clear specification of the measurement method).	Simply states number of F508del copies, none 1 or 2.	Partial
b)	Valid and Reliable Measurement of PF	Method of PF measurement is valid and reliable to limit misclassification bias (may include relevant outside sources of information on measurement properties, such as blind measurement and limited reliance on recall).	Technical method of genotyping is not given and no further information is given about genotypes assessed	No
c)	Method and Setting of PF Measurement	The method and setting of measurement of PF is the same for all study participants.	Unclear how genotyping was performed across centres and it's likely to have been carried out at different facilities.	Unsure
d)	Proportion	Adequate proportion	All reported in this study have genotyping	Unsure

	of data on PF available for analysis	(>70%) of the study sample has complete data for PF variable.	data available but unclear whether others may have been excluded who did not have this data.		
e)	Method used for missing data	Appropriate methods of imputation are used for missing PF data.	Unclear if any imputation used for genotype data recorded in the registry.	Unsure	
PF Measurement Summary		**PF is adequately measured in study participants to sufficiently limit potential bias.**			**High: assessment of link with genotype isn't the primary aim of the study and many areas unknown**
4. Outcome Measurement		**Goal: To judge the risk of bias related to the measurement of outcome (differential measurement of outcome related to the baseline level of PF).**			
a)	Definition of the Outcome	A clear definition of outcome is provided, including duration of follow-up.	Study is looking at FEV1 decline using a median 19 FEV1 observations per individual 6.8 years of follow-up	Yes	
b)	Valid and Reliable Measurement of	The method of outcome measurement used is adequately valid and reliable to	Clearly explains analysis technique using functional principal components analysis for sparse longitudinal data (FPCA). Patterns of decline grouped as early/late/middle with	Yes	

	Outcome	limit misclassification bias (may include relevant outside sources of information on measurement properties, also characteristics, such as blind measurement and confirmation of outcome with valid and reliable test).	clear definitions for each		
c)	Method and Setting of Outcome Measurement	The method and setting of outcome measurement is the same for all study participants.	FEV1 will have been measured across different centres though expected to be minimal variation in assessment	Unsure	
Outcome Measurement Summary		**Outcome of interest is adequately measured in study participants to sufficiently limit potential bias.**			**Low**
5. Study Confounding		**Goal: To judge the risk of bias due to confounding (where the effect of the PF is distorted by another factor that is related to both the PF and outcome).**			
a)	Important Confounders Measured	Important confounders including treatments are measured (key LIST variables)	Adjusts for age at baseline, at diagnosis, gender, birth cohort year, socioeconomic status and phenotypic variables	Partial	

b)	Definition of the confounding factor	Clear definitions of the important confounders measured are provided (including dose, level, and duration of exposure).	For the variables assessed	Yes
c)	Valid and Reliable Measurement of Confounders	Measurement of all important confounders is adequately valid and reliable (may include relevant outside sources of information on measurement properties, also characteristics, such as blind measurement and limited reliance on recall).	General descriptions are given though it's unclear whether all data will have been entered accurately into the registry for all patients	Unsure
d)	Method and Setting of Confounding Measurement	The method and setting of confounding measurement are the same for all study participants.	Unclear method of assessment and likely to have varied between individuals	Unsure
e)	Method used for	Appropriate methods are used if imputation	Unsure whether there may have been missing data on confounders or how this	Unsure

	missing data	is used for missing confounder data.	was managed.		
f)	Appropriate Accounting for Confounding	Important potential confounders are accounted for in the study design (for example, matching for key variables, stratification, or initial assembly of comparable groups).	No matching or stratification	No	
		Important potential confounders are accounted for in the analysis (that is, appropriate adjustment).	As above some relevant confounders are adjusted for	Partial	
Study Confounding Summary		**Important potential confounders are appropriately accounted for, limiting potential bias with respect to the relationship between PF and outcome.**			**Moderate**
6. Statistical Analysis and Reporting		**Goal: To judge the risk of bias related to the statistical analysis and presentation of results.**			
a)	Presentation of analytical	There is sufficient presentation of data to	Chi squared for overall differences in variables and logistic regression model	Yes	

	strategy	assess the adequacy of the analysis.	assessing baseline characteristics as covariates of lung function decline		
b)	Model development strategy	The strategy for model building (inclusion of variables in the statistical model) is appropriate and based on a conceptual framework or model. The selected statistical model is adequate for the design of the study.	As above	Yes	
c)	Reporting of results	There is no selective reporting of results.	None apparent	No	
Statistical Analysis and Presentation Summary		**The statistical analysis is appropriate for the design of the study, limiting potential for presentation of invalid or spurious results.**			**Low**

Summary Szczesniak et al et al 2006: Participation high; Attrition high; PF high; Outcome low; Confounding moderate; Statistical Analysis low

Table 30.13

Author and year of publication	De Boeck and Zolin 2017			
Biases	**Issues to consider for judging overall rating of "Risk of bias"**	**Study Methods & Comments**	**Rating of reporting: yes, partial, no, unsure**	**Overall rating of "Risk of bias" for domain: high, moderate, low**
1. Study Participation	**Goal: To judge the risk of selection bias: the likelihood that relationship between prognostic factor (PF) and outcome is different for participants and eligible non-participants. For registry studies, this section has considered the basis for selection for the registry (as all people with CF in the region would theoretically be eligible for analysis) as the equivalent of "participant selection" in a prospective study**			
a) Source of target population	The source population or population of interest is adequately described for key characteristics (LIST).	The study gives clear participant flow though the study but participants aren't described for key characteristics. Unclear country and centre representation from this study.	No	
b) Method used to identify population	The sampling frame and recruitment are adequately described, including methods to	This study includes patients genotyped and classified, age >6 years without transplant and FEV1 data collected in at least 2 of the 3 observation years.	Partial	

		identify the sample sufficient to limit potential bias (for example, referral patterns in health care)	Unclear representation within the registry of all with CF in Europe.	
c)	Recruitment period	Period of recruitment is adequately described	Study observation period for the registry is given (2008-10)	Yes
d)	Place of recruitment	Place of recruitment (setting and geographic location) are adequately described	European registry is said to cover 15 registries and 50 centres across 12 countries but it's not described how representative this is of countries across Europe	Partial
e)	Inclusion and exclusion criteria	Inclusion and exclusion criteria are adequately described (including explicit diagnostic criteria)	Patients genotyped and classified, age >6 years without transplant and FEV1 data collected in at least 2 years. Unclear representation of the registry as a whole from this publication	Partial
f)	Adequate study participation	There is adequate participation in the study by eligible individuals (>70%)	Final sample analysed only includes one third of those in the registry but unclear how representative the registry is of all countries and people with CF in Europe	Unsure
g)	Baseline characteristics	The baseline study sample (individuals entering the study) is	Only age range is given which was lower for those with class 1/stop codon variants and all other class combinations. Other	Partial

	adequately described for key characteristics (LIST).	characteristics are not given. Exclusions due to lack of FEV1 data or lung transplant are reported after classification and the proportions per group are equivalent reducing risk that these exclusions have excluded those with more severe genotype.		
Summary Study participation	**The study sample represents the population of interest on key characteristics, sufficient to limit potential bias for the observed relationship between the PF and outcome.**			**Moderate: limited data on European representation is given but there is clear flow-through the study**
2. Study Attrition	**Goal: To judge the risk of attrition bias (likelihood that relationship between PF and the outcome are different for completing and non-completing participants). For registry studies, we considered that this section should consider loss of participants from the analysis due to lack of available data, for example.**			
a) Proportion of baseline sample available for analysis	Response rate is adequate (proportion of study sample completing the study and providing outcome data).	Only a third of those entering the study had complete data available for analysis. 61% had genotyping and classification data but there were further exclusions due to transplant/lack of FEV1 measure	No	
b) Attempts to	Attempts to collect	Genotyping/classification was performed first	Partial	

	collect information on participants who dropped out	information on participants who dropped out of the study are described.	so this data is available prior to further exclusions. No difference in frequency of classes for those included/excluded on lung function/transplant status		
c)	Reasons and potential impact of subjects lost to follow-up	Reasons for loss to follow-up are provided.	As above the study excludes due to lack of genotyping, classification, age and inadequate FEV1 measures	Yes	
d)	Outcome and prognostic factor information on those lost to follow-up	Participants lost to follow-up are adequately described for key characteristics (LIST) with no important differences from participants.	Characteristics are not given for those excluded – Aside from as above noting no difference in exclusions according to class which reduces risk lung function/transplant exclusions may have excluded those with more severe genotype.	No	
Study Attrition Summary		**Loss to follow-up (from baseline sample to study population analysed) is not associated with key characteristics (that is, the study data adequately represent the sample) sufficient to limit potential bias to the observed relationship between PF and outcome.**			**High: due to overall high attrition rate**
3. Prognostic Factor		**Goal: To judge the risk of measurement bias related to how the PF was measured (differential measurement of PF related to the level of outcome). For studies comparing variant classes this**			

Measurement		**includes whether the system used to classify was adequately described.**		
a)	Definition of the PF	A clear definition or description of 'PF' is provided (including dose, duration of exposure, and clear specification of the measurement method).	Has analysed variant classes 1-5 using established system and has explained groupings.	Yes
b)	Valid and Reliable Measurement of PF	Method of PF measurement is valid and reliable to limit misclassification bias (may include relevant outside sources of information on measurement properties, such as blind measurement and limited reliance on recall).	Technical method of genotyping is not given. The latest documented classification system is referenced, though the specific variants grouped (including stop codon) are not given.	Partial
c)	Method and Setting of PF Measurement	The method and setting of measurement of PF is the same for all study participants.	Unclear how genotyping was performed across centres in Europe but expected to be different.	No

d)	Proportion of data on PF available for analysis	Adequate proportion (>70%) of the study sample has complete data for PF variable.	61% had genotyping data available but further exclusions due to lung transplant or few FEV1 measures further reduce number analysed.	No	
e)	Method used for missing data	Appropriate methods of imputation are used for missing PF data.	Unclear if any imputation used for genotype data recorded in the registry.	Unsure	
PF Measurement Summary		**PF is adequately measured in study participants to sufficiently limit potential bias.**			**Moderate: some detail is lacking but no clear indication of risk of bias**
4. Outcome Measurement		**Goal: To judge the risk of bias related to the measurement of outcome (differential measurement of outcome related to the baseline level of PF).**			
a)	Definition of the Outcome	A clear definition of outcome is provided, including duration of follow-up.	Study is looking at FEV1 decline across 3 consecutive years.	Yes	
b)	Valid and Reliable Measurement of Outcome	The method of outcome measurement used is adequately valid and reliable to limit misclassification bias	The study is looking at change in centiles in FEV1. It doesn't report how FEV1 was measured though this is expected to be as standard	Partial	

		(may include relevant outside sources of information on measurement properties, also characteristics, such as blind measurement and confirmation of outcome with valid and reliable test).			
c)	Method and Setting of Outcome Measurement	The method and setting of outcome measurement is the same for all study participants.	FEV1 will have been measured across different centres though expected to be minimal variation in assessment	Unsure	
Outcome Measurement Summary		**Outcome of interest is adequately measured in study participants to sufficiently limit potential bias.**			**Moderate: limited detail on assessment of FEV1**
5. Study Confounding		**Goal: To judge the risk of bias due to confounding (where the effect of the PF is distorted by another factor that is related to both the PF and outcome).**			
a)	Important Confounders Measured	Important confounders including treatments are measured (key	Adjustment for age only	No	

		LIST variables)		
b)	Definition of the confounding factor	Clear definitions of the important confounders measured are provided (including dose, level, and duration of exposure).	Only age assessed	NA
c)	Valid and Reliable Measurement of Confounders	Measurement of all important confounders is adequately valid and reliable (may include relevant outside sources of information on measurement properties, also characteristics, such as blind measurement and limited reliance on recall).	Only age assessed	NA
d)	Method and Setting of Confounding Measurement	The method and setting of confounding measurement are the same for all study participants.	Only age assessed	NA
e)	Method	Appropriate methods	Unclear	Unsure

	used for missing data	are used if imputation is used for missing confounder data.			
f)	Appropriate Accounting for Confounding	Important potential confounders are accounted for in the study design (for example, matching for key variables, stratification, or initial assembly of comparable groups).	No matching or stratification	No	
		Important potential confounders are accounted for in the analysis (that is, appropriate adjustment).	Only age is adjusted for	Yes	
Study Confounding Summary		**Important potential confounders are appropriately accounted for, limiting potential bias with respect to the relationship between PF and outcome.**			**High**
6. Statistical Analysis and Reporting		**Goal: To judge the risk of bias related to the statistical analysis and presentation of results.**			
a)	Presentation	There is sufficient	Linear regression model adjusting for age.	Partial	

	of analytical strategy	presentation of data to assess the adequacy of the analysis.			
b)	Model development strategy	The strategy for model building (inclusion of variables in the statistical model) is appropriate and based on a conceptual framework or model. The selected statistical model is adequate for the design of the study.	Limited detail on statistical analysis other than stating as above and that the Tukey–Kramer method was considered for multiple comparison adjustment of the p-values for the differences of least square means estimated from the models.	Partial	
c)	Reporting of results	There is no selective reporting of results.	None apparent	No	
Statistical Analysis and Presentation Summary		**The statistical analysis is appropriate for the design of the study, limiting potential for presentation of invalid or spurious results.**			**Moderate**
Summary de Boeck and Zolin 2017: Participation moderate; Attrition high; PF moderate; Outcome moderate; Confounding high; Statistical Analysis moderate					

Table 30.14

Author and year of publication	Dugueperoux and De Braekeleer 2005			
Biases	**Issues to consider for judging overall rating of "Risk of bias"**	**Study Methods & Comments**	**Rating of reporting: yes, partial, no, unsure**	**Overall rating of "Risk of bias" for domain: high, moderate, low**
1. Study Participation	**Goal: To judge the risk of selection bias: the likelihood that relationship between prognostic factor (PF) and outcome is different for participants and eligible non-participants. For registry studies, this section has considered the basis for selection for the registry (as all people with CF in the region would theoretically be eligible for analysis) as the equivalent of "participant selection" in a prospective study**			
a) Source of target population	The source population or population of interest is adequately described for key characteristics (LIST).	The study gives characteristics for F508del heterozygotes seen in 2000 and eligible for analysis but unclear how representative they are of the source population in France with these variants.	No	
b) Method used to identify population	The sampling frame and recruitment are adequately described, including methods to identify the sample	The study includes those registered with these variants 1992-2000 and who were seen in 2000. Says the registry covers "most of the patients seen regularly at CF care centres in	Partial	

		sufficient to limit potential bias (for example, referral patterns in health care)	France" but unclear whether there may be some differences in coverage across France.	
c)	Recruitment period	Period of recruitment is adequately described	For the registry 1992-2002 and this study analyses 2000 data	Yes
d)	Place of recruitment	Place of recruitment (setting and geographic location) are adequately described	France but unclear whether there could be some difference in distribution of care centres and coverage across the country	Partial
e)	Inclusion and exclusion criteria	Inclusion and exclusion criteria are adequately described (including explicit diagnostic criteria)	This study includes F508del heterozygotes registered and seen in 2000 with specific exclusions given.	Yes
f)	Adequate study participation	There is adequate participation in the study by eligible individuals (>70%)	16/27 (59%) registered with 3849+10kbC>T/F508del were seen during 2000. 34/61 (56%) registered with 2789+5G>A/F508del were seen during 2000. So all those that seem eligible were included but it's uncertain how representative they are of all those who have the genotype.	Unsure
g)	Baseline	The baseline study	Age, gender, age and type of diagnosis and	Yes

	characteristics	sample (individuals entering the study) is adequately described for key characteristics (LIST).	phenotypic variables given		
Summary Study participation		**The study sample represents the population of interest on key characteristics, sufficient to limit potential bias for the observed relationship between the PF and outcome.**			**Moderate: unclear whether registered patients represent all those with this variant in France but no real indication of bias**
2. Study Attrition		**Goal: To judge the risk of attrition bias (likelihood that relationship between PF and the outcome are different for completing and non-completing participants). For registry studies, we considered that this section should consider loss of participants from the analysis due to lack of available data, for example.**			
a)	Proportion of baseline sample available for analysis	Response rate is adequate (proportion of study sample completing the study and providing outcome	Of all eligible with this genotype in the 2000 assessment there are no apparent exclusions	Yes	

		data).			
b)	Attempts to collect information on participants who dropped out	Attempts to collect information on participants who dropped out of the study are described.	No information is provided for those heterozygotes not seen in 2000 though no apparent reason to suggest the year is a source of bias	Unsure	
c)	Reasons and potential impact of subjects lost to follow-up	Reasons for loss to follow-up are provided.	In the registry years only 1/39 with 3849+10kbC>T were lost to follow-up 3/39 died. Respective figures for 2789+5G>A were 4/88 and 2/88. Otherwise the sample in 2000 represents just under 60% of F508del heterozygotes those with these genotypes	Yes	
d)	Outcome and prognostic factor information on those lost to follow-up	Participants lost to follow-up are adequately described for key characteristics (LIST) with no important differences from participants.	Characteristics are not given for those included only	No	
Study Attrition Summary		**Loss to follow-up (from baseline sample to study population analysed) is not associated with key characteristics (that is, the study data adequately represent the sample) sufficient to limit potential bias to the observed relationship between PF and outcome.**			**Low: exclusions are clear and**

					no indication that those seen in 2000 should be less representative
3. Prognostic Factor Measurement		**Goal: To judge the risk of measurement bias related to how the PF was measured (differential measurement of PF related to the level of outcome). For studies comparing variant classes this includes whether the system used to classify was adequately described.**			
a)	Definition of the PF	A clear definition or description of 'PF' is provided (including dose, duration of exposure, and clear specification of the measurement method).	Has analysed specific variant	Yes	
b)	Valid and Reliable Measurement of PF	Method of PF measurement is valid and reliable to limit misclassification bias (may include relevant outside sources of information on measurement	Technical method of genotyping is not given. The study reports reconfirming the genotype for 38/39 and 82/88 who had been registered with these genotypes, but method unclear	Partial	

		properties, such as blind measurement and limited reliance on recall).			
c)	Method and Setting of PF Measurement	The method and setting of measurement of PF is the same for all study participants.	Matched homozygotes and heterozygotes came from the same centre and were reported to have been analysed using the same equipment.	Yes	
d)	Proportion of data on PF available for analysis	Adequate proportion (>70%) of the study sample has complete data for PF variable.	No apparent exclusions or lack of data (other than those with this genotype who may not have been registered)	Yes	
e)	Method used for missing data	Appropriate methods of imputation are used for missing PF data.	As above	NA	
PF Measurement Summary		**PF is adequately measured in study participants to sufficiently limit potential bias.**			**Low: the study is looking at specific genotype, has reconfirmed and patients from same centre were tested using**

					the same equipment
4. Outcome Measurement		**Goal: To judge the risk of bias related to the measurement of outcome (differential measurement of outcome related to the baseline level of PF).**			
a)	Definition of the Outcome	A clear definition of outcome is provided, including duration of follow-up.	Study is looking at variables assessed during 2000 and gives broad description of each	Yes	
b)	Valid and Reliable Measurement of Outcome	The method of outcome measurement used is adequately valid and reliable to limit misclassification bias (may include relevant outside sources of information on measurement properties, also characteristics, such as blind measurement and confirmation of outcome with valid and reliable test).	No indication that phenotypic measurement should be biased but descriptions are only general and there may be variation in how these were measured among individuals	Partial	
c)	Method and Setting of	The method and setting of outcome	As patients were matched at centres this should limit variation	Yes	

	Outcome Measurement	measurement is the same for all study participants.			
Outcome Measurement Summary		**Outcome of interest is adequately measured in study participants to sufficiently limit potential bias.**			**Moderate**
5. Study Confounding		**Goal: To judge the risk of bias due to confounding (where the effect of the PF is distorted by another factor that is related to both the PF and outcome).**			
a)	Important Confounders Measured	Important confounders including treatments are measured (key LIST variables)	Age, gender and treatment centre to account for variation in care levels	Partial	
b)	Definition of the confounding factor	Clear definitions of the important confounders measured are provided (including dose, level, and duration of exposure).	Limited applicability for age, gender, centre	NA	
c)	Valid and Reliable Measurement of Confounders	Measurement of all important confounders is adequately valid and reliable (may include relevant outside sources of information on measurement	Limited applicability for age, gender, centre	NA	

		properties, also characteristics, such as blind measurement and limited reliance on recall).		
d)	Method and Setting of Confounding Measurement	The method and setting of confounding measurement are the same for all study participants.	Participants from the same centre	Yes
e)	Method used for missing data	Appropriate methods are used if imputation is used for missing confounder data.	Unclear	Unsure
f)	Appropriate Accounting for Confounding	Important potential confounders are accounted for in the study design (for example, matching for key variables, stratification, or initial assembly of comparable groups).	Matched for age, gender and centre	Partial

	Important potential confounders are accounted for in the analysis (that is, appropriate adjustment).	Matched only	NA	
Study Confounding Summary	**Important potential confounders are appropriately accounted for, limiting potential bias with respect to the relationship between PF and outcome.**			**Moderate**
6. Statistical Analysis and Reporting	**Goal: To judge the risk of bias related to the statistical analysis and presentation of results.**			
a) Presentation of analytical strategy	There is sufficient presentation of data to assess the adequacy of the analysis.	ANOVA and Kruskal Wallis test used to compare variables with p set at 0.05	Partial	

b)	Model development strategy	The strategy for model building (inclusion of variables in the statistical model) is appropriate and based on a conceptual framework or model. The selected statistical model is adequate for the design of the study.	Study doesn't build a model, statistical comparison of variables as above. However, small numbers limit the reliability of the analysis	No	
c)	Reporting of results	There is no selective reporting of results.	None apparent	No	
Statistical Analysis and Presentation Summary		**The statistical analysis is appropriate for the design of the study, limiting potential for presentation of invalid or spurious results.**			**High: primarily due to comparison of small sample sizes**

Summary Dugueperoux and De Braekeleer 2005: Participation moderate; Attrition low; PF low; Outcome moderate; Confounding moderate; Statistical Analysis high

Table 30.15

Author and year of publication	Mackenzie et al 2017			
Biases	**Issues to consider for judging overall rating of "Risk of bias"**	**Study Methods & Comments**	**Rating of reporting: yes, partial, no, unsure**	**Overall rating of "Risk of bias" for domain: high, moderate, low**
1. Study Participation	**Goal: To judge the risk of selection bias: the likelihood that relationship between prognostic factor (PF) and outcome is different for participants and eligible non-participants. For registry studies, this section has considered the basis for selection for the registry (as all people with CF in the region would theoretically be eligible for analysis) as the equivalent of "participant selection" in a prospective study**			
a) Source of target population	The source population or population of interest is adequately described for key characteristics (LIST).	The study gives birth cohort, age at diagnosis and other phenotypic variables for the 26 P67L heterozygotes who form the sample for analysis. Uncertain whether this sample in the registry represents all those with the P67L variant in Canada	Partial	
b) Method used to identify	The sampling frame and recruitment are adequately described,	The study includes those registered with the P67L variant 1996 to 2011.Says these are all patients seen at any CF clinics across	Unsure	

	population	including methods to identify the sample sufficient to limit potential bias (for example, referral patterns in health care)	Canada and recorded in the Canadian registry. However, F508del homozygotes for comparison only came from clinics in Atlantic Canada. Uncertain how representative they are of those from other Canadian regions and so comparable to heterozygotes from across Canada	
c)	Recruitment period	Period of recruitment is adequately described	1996 to 2011	Yes
d)	Place of recruitment	Place of recruitment (setting and geographic location) are adequately described	Atlantic centres for homozygotes, heterozygotes apparently from anywhere in Canada	Yes
e)	Inclusion and exclusion criteria	Inclusion and exclusion criteria are adequately described (including explicit diagnostic criteria)	This study includes P67L heterozygotes seen 1996-2011 across Canada and F508del homozygotes from Atlantic Canada. Some uncertainty around how representative they are.	Unsure
f)	Adequate study participation	There is adequate participation in the study by eligible individuals (>70%)	This is likely to have covered all those with this genotype in Canada but not possible to say from this publication what coverage the registry has. Similarly unclear whether this represents all homozygotes in Canada	Unsure
g)	Baseline	The baseline study	As above birth cohort, age at diagnosis and	Partial

characteristics	sample (individuals entering the study) is adequately described for key characteristics (LIST).	other phenotypic variables. No data on gender or ethnicity.		
Summary Study participation	**The study sample represents the population of interest on key characteristics, sufficient to limit potential bias for the observed relationship between the PF and outcome.**			**High: primarily on basis of uncertainty around whether homozygotes representative and comparable**
2. Study Attrition	**Goal: To judge the risk of attrition bias (likelihood that relationship between PF and the outcome are different for completing and non-completing participants). For registry studies, we considered that this section should consider loss of participants from the analysis due to lack of available data, for example.**			
a) Proportion of baseline sample available for analysis	Response rate is adequate (proportion of study sample completing the study and providing outcome data).	For all registered with this variant and comparison homozygotes, there are no exclusions for the assessment of age at diagnosis and pancreatic status. NB these are the only variables analysed as there was minimal data for others.	Yes	

b)	Attempts to collect information on participants who dropped out	Attempts to collect information on participants who dropped out of the study are described.	No apparent loss for those registered.	NA	
c)	Reasons and potential impact of subjects lost to follow-up	Reasons for loss to follow-up are provided.	No apparent loss for those registered.	NA	
d)	Outcome and prognostic factor information on those lost to follow-up	Participants lost to follow-up are adequately described for key characteristics (LIST) with no important differences from participants.	No apparent loss for those registered.	NA	
Study Attrition Summary		**Loss to follow-up (from baseline sample to study population analysed) is not associated with key characteristics (that is, the study data adequately represent the sample) sufficient to limit potential bias to the observed relationship between PF and outcome.**			**Low: those identified have been assessed for age and**

					pancreatic status
3. Prognostic Factor Measurement		**Goal: To judge the risk of measurement bias related to how the PF was measured (differential measurement of PF related to the level of outcome). For studies comparing variant classes this includes whether the system used to classify was adequately described.**			
a)	Definition of the PF	A clear definition or description of 'PF' is provided (including dose, duration of exposure, and clear specification of the measurement method).	Has analysed specific variant	Yes	
b)	Valid and Reliable Measurement of PF	Method of PF measurement is valid and reliable to limit misclassification bias (may include relevant outside sources of information on measurement properties, such as blind measurement and limited reliance on recall).	No mention is given to the method of genotyping and it may have differed across centres.	Unsure	

c)	Method and Setting of PF Measurement	The method and setting of measurement of PF is the same for all study participants.	Unlikely to be the same across centres.	Unsure	
d)	Proportion of data on PF available for analysis	Adequate proportion (>70%) of the study sample has complete data for PF variable.	No apparent exclusions or lack of data (other than those with this genotype who may not have been registered)	Yes	
e)	Method used for missing data	Appropriate methods of imputation are used for missing PF data.	As above	NA	
PF Measurement Summary		**PF is adequately measured in study participants to sufficiently limit potential bias.**			**Moderate: study is looking at specific variant, no real indication of bias but little information is given**
4. Outcome Measurement		**Goal: To judge the risk of bias related to the measurement of outcome (differential measurement of outcome related to the baseline level of PF).**			
a)	Definition of	A clear definition of	Age at diagnosis and pancreatic status	Partial	

	the Outcome	outcome is provided, including duration of follow-up.	assessed here – unclear whether the latter was just a one-off status measure		
b)	Valid and Reliable Measurement of Outcome	The method of outcome measurement used is adequately valid and reliable to limit misclassification bias (may include relevant outside sources of information on measurement properties, also characteristics, such as blind measurement and confirmation of outcome with valid and reliable test).	Pancreatic status was taken by the recording of ERT as is standard across registry studies but unclear whether there could be error in this	Partial	
c)	Method and Setting of Outcome Measurement	The method and setting of outcome measurement is the same for all study participants.	Likely to have varied across centres in Canada	Unsure	
Outcome Measurement Summary		**Outcome of interest is adequately measured in study participants to sufficiently limit potential bias.**			**Moderate**

5. Study Confounding		**Goal: To judge the risk of bias due to confounding (where the effect of the PF is distorted by another factor that is related to both the PF and outcome).**		
a)	Important Confounders Measured	Important confounders including treatments are measured (key LIST variables)	No adjustment for confounders	No
b)	Definition of the confounding factor	Clear definitions of the important confounders measured are provided (including dose, level, and duration of exposure).	No adjustment for confounders	NA
c)	Valid and Reliable Measurement of Confounders	Measurement of all important confounders is adequately valid and reliable (may include relevant outside sources of information on measurement properties, also characteristics, such as blind measurement and limited reliance on recall).	No adjustment for confounders	NA
d)	Method and	The method and	No adjustment for confounders	NA

	Setting of Confounding Measurement	setting of confounding measurement are the same for all study participants.			
e)	Method used for missing data	Appropriate methods are used if imputation is used for missing confounder data.	No adjustment for confounders	NA	
f)	Appropriate Accounting for Confounding	Important potential confounders are accounted for in the study design (for example, matching for key variables, stratification, or initial assembly of comparable groups).	No matching for age, gender and centre	No	
		Important potential confounders are accounted for in the analysis (that is, appropriate adjustment).	No adjustment for confounders	NA	
Study Confounding		**Important potential confounders are appropriately accounted for, limiting potential bias with respect to the relationship between PF and outcome.**			**High: particularly**

Summary				**given specific location of care centre for homozygotes**
6. Statistical Analysis and Reporting	**Goal: To judge the risk of bias related to the statistical analysis and presentation of results.**			
a) Presentation of analytical strategy	There is sufficient presentation of data to assess the adequacy of the analysis.	t test and Mann-Whitney used to compare respectively age at diagnosis and pancreatic status.	Partial	
b) Model development strategy	The strategy for model building (inclusion of variables in the statistical model) is appropriate and based on a conceptual framework or model. The selected statistical model is adequate for the design of the study.	Study doesn't build a model, statistical comparison of variables as above. However, small numbers limit the reliability of the analysis	No	
c) Reporting of results	There is no selective reporting of results.	None apparent	No	

Statistical Analysis and Presentation Summary	**The statistical analysis is appropriate for the design of the study, limiting potential for presentation of invalid or spurious results.**	**High: primarily due to comparison of small sample sizes**
Summary Mackenzie et al 2017: Participation high; Attrition low; PF moderate; Outcome moderate; Confounding high; Statistical Analysis high		

QUADAS quality assessment of test accuracy study

Table 31: Massie et al 2009[42] (Criteria 4 and 8)

Domain	Risk of Bias	Notes
Domain I: Patient selection		
Consecutive or random sample of population enrolled?	Unclear	Pay-for service so though offered to all in region may not be representative of all
Case-control design avoided?	Low	
Inappropriate exclusions avoided?	Unclear	As above, not exclusions but charge for testing may influence representation
Domain II: Index test		
Index test results interpreted without knowledge of reference standard results?	NA	Screen test results used to guide decision for antenatal diagnostic test
Threshold pre-specified?	NA	Panel of variants given
Domain III: Reference standard		
Reference standard likely to correctly classify condition?	High	Test should have high analytical validity to identify panel of variants being tested but screen negatives may have other CF variants
Reference standard results interpreted without knowledge of index test results?	NA	Screen test results used to guide decision for antenatal diagnostic test
Domain IV: Test strategy flow and timing		
Appropriate interval between index test and reference standard?	NA	
Did all participants receive same reference standard?	Low	Same screening and diagnostic test

All patients included in analysis?	High	No follow-up of screen negatives, variants/CF in the fetus/newborn only determined for screen positive couples
Domain V: Applicability		
Applicable to UK screening population of interest?	High	Variant frequency expected to differ between populations
Applicable to UK screening test of interest?	High	Variant panel is likely to differ from what would be used in the UK Also included pre-conception screening

Appendix 5 – Full discussion of quality appraisal for genotype-phenotype association studies

Table 8. Summary QUIPS assessments

Study	Summary risk of bias by domain					
	Participation	Attrition	Genotype measure	Phenotype measure	Confounding	Statistical analysis
McKone et al 2006[25]	moderate	high	moderate	moderate	moderate	low
McKone et al 2003[26]	low	moderate	moderate	moderate	high	low
Lai et al 2004[27]	moderate	high	high	high	high	moderate
O'Connor et al 2002[28]	moderate	high	high	high	moderate	low
Simmonds et al 2009[29]	high	high	high	moderate	high	high
Badet et al 2004[30]	high	moderate	high	low	high	high
Koch et al 2001[31]	high	moderate	moderate	high	high	N/A
Dewulf et al 2015[24]	moderate	moderate	low	moderate	moderate (treatment), high (other)	low
Green et al 2010[32]	high	moderate	moderate	low (infection), high (other)	moderate (infection), high (other)	low
Radtke et al 2017[33]	high	high	moderate	moderate	high	moderate
CF G-P Consortium 1993[34]	high	high	moderate	moderate	moderate	moderate
Szczesniak et al 2017[36]	high	high	high	low	moderate	low
de Boeck and Zolin 2017[37]	moderate	high	moderate	moderate	high	moderate
Dugueperoux de Braekeleer	moderate	low	low	moderate	moderate	high

Study	Summary risk of bias by domain					
	Participation	Attrition	Genotype measure	Phenotype measure	Confounding	Statistical analysis
2005[35]						
Mackenzie et al 2017[23]	high	low	moderate	moderate	high	high

Study participation and attrition

The main strength of studies is that by analysing data from national CF registries or international consortiums they had data for several thousand participants. This should give increased power for detecting differences in phenotype according to genotype. However, there are inherent limitations when using collective data in national registries.

The participation component of QUIPS was assessed based on how well participants represented the general population with CF who would theoretically be eligible for analysis. CF registries would be expected to include the vast majority of people with CF from the countries or regions studied. However, most studies did not report how representative the registry was of all people with CF in that country. Most studies also did not clarify by what process people are reported to the registries or how regularly their clinical data is entered.

Studies mostly scored moderate risk of bias for participation because they did not clarify the national coverage that the registry gives for all people with CF. The exception of the low risk assessment for Mckone et al[26] is because this study has given more information on the national coverage of the registry, including the proportion of people with CF represented. It also gave comparison characteristics for the registry population not covered by the study. It is accepted that most national registries would be expected to have similarly high coverage of all people with CF, but this was not assumed. However, because neither was there indication that the study or registry gave biased or incomplete representation, these studies were rated moderate rather than high risk of bias.

Studies rated to have high risk of bias mostly had specific representation issues, as follows:

Badet et al[30] aimed to look at people within the CF registry surviving to over 30 years. However, they specifically excluded people diagnosed above the age of 5 years. This could make genotype comparison

unreliable as it may have excluded survivors with milder genotype who were diagnosed in older childhood/early adulthood. Simmonds et al[29] (also looking at older survivors) only included those from a single centre, with comparison to the full registry. People treated at this centre may differ from older people in the full UK registry. Mackenzie et al[23] had similar issue in identified P67L heterozygotes from across Canada but comparing them with F508del homozygotes from only one geographic region in Canada.

The study of exercise capacity by the Exercise Working Group of the European CF Society[33] only included those aged above 8 years who had completed a maximal effort during exercise performance, so may exclude those with severe disease. Whereas Green et al[32] required participants to have a surviving sibling also with CF. Two European studies also had risk of poor representation in covering less than 50% of people with CF from across the eligible countries.[31, 34]

The most common reason for attrition in nearly all studies was not participant drop-out as such, but lack of genotyping or classification for the available registry population. Most registry studies had genotyping (and classification data where relevant) available for only between 50%[25, 28] and 75% of the full registry cohort.[24, 31] Some studies applying further inclusion criteria, such as a minimum age or requiring follow-up assessments, had data for far smaller subsamples of 10-30% of the full registry. While these inclusion criteria are understandable, the small proportions of the registries included may mean results are less representative of the population with CF.

This lack of genotyping or classification was the reason for high risk of bias related to attrition in nearly all studies, except for a few with better coverage of their studied population. This included studies of specific variants which appear to have included all individuals with that genotype in the registry with no apparent exclusions.[23, 35]

Studies varied in whether or not they described characteristics for both the full potential registry cohort and those genotyped (or with classification) and so available for analysis. Optimally some studies, such as McKone et al (2003)[26] have listed characteristics for those genotyped/analysed and those not analysed (though without statistical comparison). Others mostly gave characteristics for the full registry cohort

or for those genotyped, but not both. As such it is difficult to know whether there may be important differences between those who have been genotyped or not, which could bias analyses.

If non-inclusion in study registries, or lack of genotyping, was random then this may be less of a problem. However, if it was selective then this could mean that studies may not represent all people with CF. It could be, for example, that patients in certain geographic locations or socioeconomic groups may be less likely to be genotyped or included in registries. Likewise, people with more severe disease manifestations with regular clinic attendance may be more likely to be genotyped than those with mild disease manifestations. Alternatively there could be survivor bias, where those living longer and with repeated follow-up assessments are more likely to be registered and genotyped. In support of this latter possibility, case-controls looking at survival could access genotyping data for over 80-90% of older patients[29, 30] which was higher than the typical genotyping rate seen for the complete registries.

Genotype assessment

Common to all studies was a lack of technical information on genotyping. Methods used may have differed between the individual centres providing data to the registry/consortium, and over time within individual centres.

For studies that classified variants by functional effect, there are two areas of limitations: potential for misclassification and absent classification. Numerous studies published during the 1990s and beyond began to describe the functional effect of different variants.[18-22] A definitive list was not identified by this review, but the list most recently updated by De Boeck et al (2014)[44] which built on that previously reported by McKone et al,[25] was used as standard in our classification of variants (as Table 13, question 3). However, there has been variation in how individual variants are classified by different research groups and as a result there are some discrepancies between studies. For example, G85E was initially classified as mild class 4[26, 27, 31] but has more recently been reassigned to severe class 2.[25, 37] Meanwhile Lai et al[27] differed in classifying 2789+5G>A and A455E as respectively class 1 and 3, when these are more commonly accepted as mild class 5 variants.[25, 44] Lai et al[27] report this as being consistent with the classification system originally developed by Welsh and Smith.[20] However, the Welsh and Smith publication only gives selected examples and does not name these

variants, specifically. These discrepancies likely reflect the challenges in classifying variants into mutually exclusive groups.

The above individual variants are not particularly common, and so would have contributed smaller numbers to the cohort as a whole. However, there is the possibility that variability in grouping could affect the reliability and comparability of analyses by class.

Not all variants have been classified in the functional classification system (or may not have been at the time of earlier studies). This is another limitation of analysis by functional classification. However, this issue does not so much relate to bias in measurement of the prognostic factor, but to attrition due to loss of data for individuals with unclassified variants (as discussed). It is also relevant to the utility of these analyses for predicting the outcome for all individuals with CF.

Aside from inconsistencies in classification, the other common reason for scoring high risk of bias in this domain applied to studies comparing F508del homozygotes with F508del heterozygotes or heterozygotes not carrying the F508del variant. Several of these studies gave incomplete or no information on the secondary variant, or grouped them according to whether the second variant was “known” or “unknown”. [28-30, 36] Therefore the specific genotype assessment was unknown which limits interpretation as these results relate to heterogeneous groups.

Studies scoring lower risk were those with high rates of genotyping for the available population, single centres performing genotyping or efforts to reconfirm the genotype.[24, 35]

Phenotype assessment
The included registry studies were clear in reporting the phenotypic variables that they have assessed. However, they rely upon pre-collected and pre-recorded clinical data, usually collected across multiple centres. There may be variation across centres in who assessed outcomes, by what method, and how outcomes were defined. It is also mostly unclear how this data was entered into the registries, and how often it was reported for individuals.

Survival assessments were conducted in variable ways. Some studies identified people alive beyond specific age cut-offs of 30 or 40 years,

while one study[26] calculated standardised mortality rates. Others provided more limited information in looking at risk of death or survival, but not indicating what life expectancy may be. For example Lai et al[27] looked at dichotomous variables of “longer” and “shorter” survival without further defining this.

Neither did studies describe how they identified patient deaths. There is no mention of accessing medical records or mortality registries and it is expected that deaths have been recorded in CF registries. But it is difficult to judge whether records are complete and up-to-date. Survival studies also differed in whether they counted transplant receipt as mortality[26, 29] (based on the assumption that the patient would have died without transplant) or whether they did not state their approach to this issue.[25, 27, 28, 30]

Looking at other phenotypic variables, lung function may be expected to be recorded in a relatively standardised way by spirometry across centres. Similarly age at diagnosis may be expected to be consistently interpreted across centres, though would likely include highly variable presentations (for example, clinical symptoms, family history or screening). *P. aeruginosa* colonisation was most often assessed by looking at positive sputum cultures over a one-year period. But it is unclear how consistently this may have been measured and entered into registries for individuals across centres.

Pancreatic insufficiency has commonly been defined as use of enzyme replacement therapy (ERT). This may not be a precise indicator and could indicate varying degrees of insufficiency. The type of ERT, dose, frequency and duration of prescription may vary considerably between individuals across centres and between studies.

A couple of studies carried out prospective assessments looking at decline in lung function over consecutive years or assessments.[36, 37] Others described only analysing individuals with >1 follow-up assessment.[25, 27, 53] However, most studies do not clarify whether phenotypic measures have been averaged across multiple assessments for each individual or whether they are just one-off measures. Therefore though many studies are retrospective cohorts assessing set years within the registry, genotype-phenotype assessments could be effectively cross sectional, for example, looking at single data entries recorded at registry

entry. This may be suitable for fixed measures such as age at diagnosis, but lung function, bacterial colonisation, pancreatic sufficiency and nutritional status could all vary over time.

For these reasons most studies have assess moderate-high risk of bias for phenotype measure. The few studies with low risk have looked at individual clearly defined outcomes (current age,[30] annual FEV1 decline[36] and infection by different criteria[32]) and have lower potential for bias.

Confounding

There was generally a high risk of bias related to confounding across studies. Key confounders considered of relevance were age, gender, ethnicity, age and method of diagnosis, and treatment received – or study centre or birth cohort as proxy measures for this.

Few studies adjusted for confounders and those that did varied in those assessed. High risk of bias reflected studies that gave no adjustment for confounders. Moderate risk was applied for studies that attempted to adjustment for some, but not all, relevant factors.

No study adjusted for treatment, per se. However, some studies made attempt to account for geographic or temporal differences in care availability. McKone et al (2006),[25] probably the most informative study for predicting survival outlook from genotype, adjusted for age and other phenotypic variables in addition to cohort year and size of treatment centre. These latter variables may be considered rough proxy indicators for care received. Two studies similarly matched homozygotes and heterozygotes from the same care centre[34, 35] or adjusted for cohort year[36] which may allow some consistency in care received. Aside from this no study adjusted for treatment.

Even had there been greater adjustment for treatment, there may still be limited applicability to CF care in the UK today. While the underlying relationship between genotype and phenotype may be expected to stay the same over time, improvements in supportive care and the availability of disease-specific treatment could alter disease course for many genotypes. Despite being representative of Western countries, the vast majority of studies looked at cohorts from around 20-30 years ago. Survival outlook has improved since then and disease manifestations may be better controlled. The disease-specific treatment ivacaftor has

only been approved within the last few years. As this acts by correcting the underlying CFTR gating problem in individuals with class 3 variants, it is likely to vastly improve outlook for these severe variants.

O'Connor et al[28] identified variables to adjust for in analysis of CF mortality. They adjusted for gender, age and type of presentation in addition to ethnicity and socioeconomic status. Green et al[32] and Szczesniak et al[36] also respectively accounted for ethnicity and socioeconomic status. No study adjusted for geographic region or country (relevant to international[33] or European studies[31, 37]). Ethnicity and environmental background may both influence genotype prevalence and disease outlook.

Other moderate risk studies included adjustment for age, gender and relevant phenotypic variables (for example, lung function and number of cultures performed in assessment of infection[32]).

The uncertain newborn screening context is one variable that could influence analyses, particularly that for age at diagnosis. No study adjusted for screening. Some studies report that newborn screening was not performed[23, 24, 28, 35] but for others this is unclear.[26, 34] Most study periods pre-date the Millennium (with birth of included cohorts even earlier) so would likely have been conducted prior to the widespread implementation of newborn screening. But there could be variability within US states and across European countries in the timing of introduction.

Statistical analyses

The statistical approach used varied between groups. Some studies were designed with the objective of developing Cox proportional hazards models to look at whether genotype can predict survival or other outcomes.[26-28, 32]

Other studies conducted regression analyses or used variable statistical tests to compare characteristics between groups. Koch et al[31] differed in comparing means and the overlap of 95% confidence intervals between groups without statistical comparison.

Most studies had reasonable sample size when comparing broad categories of severe/mild class or homozygotes/heterozygotes. However, several analyses became small when looking at rarer genotypes[23, 26, 34, 35]

or when comparing subgroups of cases with longer survival.[30, 36] These small samples may mean that the results are less representative of the population with these genotypes/phenotypes as a whole.

The p value threshold for significance also varied between studies, with some studies adjusting the level required for significance due to multiple testing, but others not. For example, McKone et al[26] set $p<0.01$ as the threshold for significance for their survival analyses and $p<0.001$ for other phenotypic variables. Most studies reported significance at $p<0.001$ but did not clarify what threshold had been set for significance. Some studies specifically stated that significance was taken at the standard level $p<0.05$.[30, 35]

Appendix 6 – 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 31.

Table 32. 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.	6
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 Objectives – What are the questions the current evidence summary intends to answer? – statement of the key questions for the current evidence summary,	12 16

		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.	18
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*.	18-22
2.3	Appraisal for quality/risk of bias tool	Details of tool/checklist used to assess quality, e.g. QUADAS 2, CASP, SIGN, AMSTAR.	23
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.	18-20
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.	93-104
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.	18, 24, 32, 69, 83
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	Study level reporting: 117 Quality assessment: 158

	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.	
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.	25, 34, 72, 83
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.	26, 54, 72, 84
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'?	30, 66, 81, 87
6.	REVIEW SUMMARY		
6.1	Conclusions and implications for policy	Do findings indicate whether screening should be recommended? Is further work warranted? Are there gaps in the evidence highlighted by the review?	90
6.2	Limitations	Discuss limitations of the available evidence and of the	92

review methodology if relevant.

References

1. Cystic Fibrosis T. UK cystic fibrosis registry annual data report 2016. London: Cystic Fibrosis Trust, 2017. Available from: https://www.cysticfibrosis.org.uk/~/media/documents/the-work-we-do/uk-cf-registry/uk-cf-registry-annual-data-report-2016.ashx?la=en.
2. Cystic Fibrosis T. UK CF registry annual data report 2008. London: Cystic Fibrosis Trust. Available from: https://www.cysticfibrosis.org.uk/~/media/documents/the-work-we-do/uk-cf-registry/2008-registry-annual-data-report.ashx?la=en.
3. Murray J, Cuckle H, Taylor G, et al. Screening for cystic fibrosis. Health Technol Assess. 1999;3:i-iv, 1-104.
4. Wald NJ, Morris JK, Rodeck CH, et al. Cystic fibrosis: Selecting the prenatal screening strategy of choice. Prenatal Diagnosis. 2003;23(6):474-83.
5. Brennan ML, Schrijver I. Cystic Fibrosis: A Review of Associated Phenotypes, Use of Molecular Diagnostic Approaches, Genetic Characteristics, Progress, and Dilemmas. Journal of Molecular Diagnostics. 2016;18(1):3-14.
6. Dorfman R, Zielenski J. Genotype-phenotype correlations in cystic fibrosis. In: Bush A, Alton EWFW, Davies JC, et al., editors. Cystic Fibrosis in the 21st Century Prog Respir Res. 34. Basel: Karger; 2006. p. 61-8.
7. Cystic fibrosis mutation database. 2011.
8. CFTR2. Clinical and functional translation of CFTR [Internet]. Baltimore (MD): US CF Foundation. Available from: https://www.cftr2.org/.
9. Fanen P, Wohlhuter-Haddad A, Hinzpeter A. Genetics of cystic fibrosis: CFTR mutation classifications toward genotype-based CF therapies. International Journal of Biochemistry and Cell Biology. 2014;52:94-102.
10. Zielenski J. Genotype and phenotype in cystic fibrosis. Respiration. 2000;67(2):117-33.
11. Committee UKNS. A laboratory guide to newborn screening in the UK for cystic fibrosis. Public Health England, 2014. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/397726/Cystic_Fibrosis_Lab_Guide_February_2014_v1.0_12_.pdf.
12. Cystic Fibrosis T. UK cystic fibrosis registry annual data report 2012. London: Cystic Fibrosis Trust, 2013. Available from: https://www.cysticfibrosis.org.uk/~/media/documents/the-work-we-do/uk-cf-registry/2012-uk-cf-registry-national-data-report.ashx?la=en.
13. Hoo ZH, Wildman MJ, Teare MD. Exploration of the impact of 'mild phenotypes' on median age at death in the UK CF registry. Respiratory Medicine. 2014;108(5):716-21.
14. Ons. Time series: United Kingdom population mid-year estimate [Internet]. Newport: Office for National Statistics. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/timeseries/ukpop/pop.
15. Ons. Births in England and Wales: 2016: live births, stillbirths, and the intensity of childbearing measured by the total fertility rate [Internet]. Newport: Office for National Statistics. Available from: https://www.ons.gov.uk/generator?uri=/peoplepopulationandcommunity/birthsdeathsandmarriages/livebirths/bulletins/birthsummarytablesenglandandwales/2016/da482e8a&format=xls.
16. Nrs. 2016 births, deaths and other vital events: preliminary annual figures [Internet]. Edinburgh: National Records of Scotland. Available from:

https://www.nrscotland.gov.uk/files//statistics/births-marriages-deaths-preliminary/2016/prelim-16-alltabs-corrected.xlsx.
17. Nisra. Live births, 1887 to 2015 [Internet]. Belfast: Northern Ireland Statistics Research Agency. Available from: https://www.nisra.gov.uk/sites/nisra.gov.uk/files/publications/live_births1887-2015.xls.
18. Kerem BS, Rommens JM, Buchanan JA, et al. Identification of the cystic fibrosis gene: Genetic analysis. Science. 1989;245(4922):1073-80.
19. Tsui LC. The spectrum of cystic fibrosis mutations. Trends in Genetics. 1992;8(11):392-8.
20. Welsh MJ, Smith AE. Molecular mechanisms of CFTR chloride channel dysfunction in cystic fibrosis. Cell. 1993;73(7):1251-4.
21. Wilschanski M, Zielenski J, Markiewicz D, et al. Correlation of sweat chloride concentration with classes of the cystic fibrosis transmembrane conductance regulator gene mutations. Journal of Pediatrics. 1995;127(5):705-10.
22. Zielenski J, Tsui LC. Cystic fibrosis: Genotypic and phenotypic variations. Annual Review of Genetics1995. p. 777-807.
23. MacKenzie IER, Paquette V, Gosse F, et al. Modeling cystic fibrosis disease progression in patients with the rare CFTR mutation P67L. Journal of cystic fibrosis. 2017;16(3):335-41.
24. Dewulf J, Vermeulen F, Wanyama S, et al. Treatment burden in patients with at least one class IV or v CFTR mutation. Pediatric Pulmonology. 2015;50(12):1230-6.
25. McKone EF, Goss CH, Aitken ML. CFTR genotype as a predictor of prognosis in cystic fibrosis. Chest. 2006;130(5):1441-7.
26. McKone EF, Emerson SS, Edwards KL, et al. Effect of genotype on phenotype and mortality in cystic fibrosis: A retrospective cohort study. Lancet. 2003;361(9370):1671-6.
27. Lai HJ, Cheng Y, Cho H, et al. Association between Initial Disease Presentation, Lung Disease Outcomes, and Survival in Patients with Cystic Fibrosis. American Journal of Epidemiology. 2004;159(6):537-46.
28. O'Connor GT, Quinton HB, Kahn R, et al. Case-mix adjustment for evaluation of mortality in cystic fibrosis. Pediatric Pulmonology. 2002;33(2):99-105.
29. Simmonds NJ, Cullinan P, Hodson ME. Growing old with cystic fibrosis - The characteristics of long-term survivors of cystic fibrosis. Respiratory Medicine. 2009;103(4):629-35.
30. Badet F, Bellis G, De Braekeleer M, et al. Phenotype and genotype of French cystic fibrosis patients with long survival and follow-up. European Journal of Internal Medicine. 2004;15(4):238-41.
31. Koch C, Cuppens H, Rainisio M, et al. European epidemiologic registry of cystic fibrosis (ERCF): Comparison of major disease manifestations between patients with different classes of mutations. Pediatric Pulmonology. 2001;31(1):1-12.
32. Green DM, McDougal KE, Blackman SM, et al. Mutations that permit residual CFTR function delay acquisition of multiple respiratory pathogens in CF patients. Respiratory research. 2010;11.
33. Radtke T, Hebestreit H, Gallati S, et al. CFTR genotype and maximal exercise capacity in cystic fibrosis a cross-sectional study. Annals of the american thoracic society. 2018;15(2):209-16.
34. The Cystic Fibrosis Genotype-Phenotype Consortium. Correlation between Genotype and Phenotype in Patients with Cystic Fibrosis. New England Journal of Medicine. 1993;329(18):1308-13.

35. Duguépéroux I, De Braekeleer M. The CFTR 3849+10kbC->T and 2789+5G->A alleles are associated with a mild CF phenotype. European Respiratory Journal. 2005;25(3):468-73.
36. Szczesniak RD, Li D, Su W, et al. Phenotypes of rapid cystic fibrosis lung disease progression during adolescence and young adulthood. American journal of respiratory and critical care medicine. 2017;196(4):471-8.
37. De Boeck K, Zolin A. Year to year change in FEV1 in patients with cystic fibrosis and different mutation classes. Journal of cystic fibrosis. 2017;16(2):239-45.
38. Gan KH, Geus WP, Bakker W, et al. Genetic and clinical features of patients with cystic fibrosis diagnosed after the age of 16 years. Thorax. 1995;50(12):1301-4.
39. Johansen HK, Nir M, Koch C, et al. Severity of cystic fibrosis in patients homozygous and heterozygous for ΔF508 mutation. The Lancet. 1991;337(8742):631-4.
40. Kerem E, Corey M, Kerem BS, et al. The relation between genotype and phenotype in cystic fibrosis - Analysis of the most common mutation (ΔF508). New England Journal of Medicine. 1990;323(22):1517-22.
41. Lester LA, Kraut J, Lloyd-Still J, et al. ΔF508 genotype does not predict disease severity in an ethnically diverse cystic fibrosis population. Pediatrics. 1994;93(1):114-8.
42. Massie J, Petrou V, Forbes R, et al. Population-based carrier screening for cystic fibrosis in Victoria: The first three years experience: Original Article. Australian and New Zealand Journal of Obstetrics and Gynaecology. 2009;49(5):484-9.
43. Network NUGT. Cystic fibrosis [Internet]. London: NHS UK Genetic Testing Network. Available from: https://ukgtn.nhs.uk/find-a-test/search-by-disorder-gene/cystic-fibrosis-1/#c3594.
44. De Boeck K, Zolin A, Cuppens H, et al. The relative frequency of CFTR mutation classes in European patients with cystic fibrosis. Journal of cystic fibrosis. 2014;13(4):403-9.
45. Pisnoli L, O'Connor A, Goldsmith L, et al. Impact of fetal or child loss on parents' perceptions of non-invasive prenatal diagnosis for autosomal recessive conditions. Midwifery. 2016;34:105-10.
46. Wright KF, Bryant LD, Morley S, et al. Presenting life with cystic fibrosis: a Q-methodological approach to developing balanced, experience-based prenatal screening information. Health expectations : an international journal of public participation in health care and health policy. 2015;18(5):1349-62.
47. Ioannou L, Massie J, Collins V, et al. Population-based genetic screening for cystic fibrosis: attitudes and outcomes. Public Health Genomics2010. p. 449-56.
48. Ioannou L, Delatycki MB, Massie J, et al. "Suddenly Having two Positive People who are Carriers is a Whole New Thing" - Experiences of Couples Both Identified as Carriers of Cystic Fibrosis Through a Population-Based Carrier Screening Program in Australia. Journal of Genetic Counseling. 2015;24(6):987-1000.
49. Janssens S, Chokoshvilli D, Binst C, et al. Attitudes of cystic fibrosis patients and parents toward carrier screening and related reproductive issues. European Journal of Human Genetics. 2016;24(4):506-12.
50. Boardman FK, Hale R. How do genetically disabled adults view selective reproduction? Impairment, identity, and genetic screening. Molecular genetics & genomic medicine. 2018.
51. Ahmed N, Corey M, Forstner G, et al. Molecular consequences of cystic fibrosis transmembrane regulator (CFTR) gene mutations in the exocrine pancreas. Gut. 2003;52(8):1159-64.
52. Sebro R, Levy H, Schneck K, et al. Cystic fibrosis mutations for p.F508del compound heterozygotes predict sweat chloride levels and pancreatic sufficiency. Clinical Genetics. 2012;82(6):546-51.

53. Dray X, Kanaan R, Bienvenu T, et al. Malnutrition in adults with cystic fibrosis. European Journal of Clinical Nutrition. 2005;59(1):152-4.
54. Corey M, Edwards L, Levison H, et al. Longitudinal analysis of pulmonary function decline in patients with cystic fibrosis. Journal of Pediatrics. 1997;131(6):809-14.
55. Courtney JM, Bradley J, McCaughan J, et al. Predictors of mortality in adults with cystic fibrosis. Pediatric Pulmonology. 2007;42(6):525-32.
56. Borgo G, Mastella G, Gasparini P, et al. Pancreatic function and gene deletion F508 in cystic fibrosis. Journal of Medical Genetics. 1990;27(11):665-9.
57. Borgo G, Gasparini P, Bonizzato A, et al. Cystic fibrosis: The ΔF508 mutation does not lead to an exceptionally severe phenotype. A cohort study. European Journal of Pediatrics. 1993;152(12):1006-11.
58. Sanders DB, Li Z, Laxova A, et al. Risk factors for the progression of cystic fibrosis lung disease throughout childhood. Annals of the american thoracic society. 2014;11(1):63-72.
59. Geborek A, Hjelte L. Association between genotype and pulmonary phenotype in cystic fibrosis patients with severe mutations. Journal of cystic fibrosis. 2011;10(3):187-92.
60. Kristidis P, Bozon D, Corey M, et al. Genetic determination of exocrine pancreatic function in cystic fibrosis. American Journal of Human Genetics. 1992;50(6):1178-84.