April 2014, Volume 1, Number 2

A systematic review to evaluate the effectiveness of enzyme replacement therapy for lysosomal storage disorders in comparison to the treatment of similar diseases with higher prevalence

Shona Lang[1], Nigel Armstrong[1], Caro Noake[1], Jos Kleijnen[2]

[1]Kleijnen Systematic Reviews Ltd., Unit 6, Escrick Business Park, Riccall Road, Escrick, York, YO19 6FD, [2]School for Public Health and Primary Care (CAPHRI), Maastricht University, Maastricht, Netherlands

ABSTRACT

Significant challenges exist when submitting orphan drugs for reimbursement in a health care system. An assessment of value for money is undertaken by comparing the costs and clinical effectiveness of the new technology to that of an existing treatment. Orphan drugs rarely have standard treatments, with which to contrast a new technology. We aimed to compare the clinical effectiveness of enzyme replacement therapy (ERT) for rare diseases, to that of recommended drugs for comparable but non-rare diseases (lifelong treatment, shortened life expectancy due to the disease), using number needed to benefit (NNTB).

A systematic review was performed to identify randomised controlled trials (RCTs) of ERT treatments in three rare diseases; Fabry Disease, Hunter Syndrome (MPS II) and Gaucher Disease Type 1. MEDLINE, MEDLINE In-Process Citations and Daily Update, EMBASE, Cochrane Database of Systematic Reviews, Cochrane Central Register of Controlled Trials, Database of Abstracts of Reviews of Effects, Health Technology Assessment Database, ClinicalTrials.gov, metaRegister of Controlled Trials and the WHO International Clinical Trials Registry Platform were searched from inception to August 2012. We identified comparator studies from the NICE (UK) or IQWiG/G-BA (Germany) websites to 2012. For each study NNTB analyses were performed for the primary outcome of all (recommended) drug doses at all follow-up times.

Ten ERT RCTs were identified (8 in Fabry Disease, 2 in Hunter Syndrome, 0 in Gaucher Disease). Eleven analyses were performed for ERTs; NNTB calculated from the mean absolute risk difference ranged from 1.4 to 17.2 (median = 2.7). Seven comparator disease studies were identified in multiple sclerosis, rheumatoid arthritis, type 2 diabetes mellitus, peripheral arterial disease and Alzheimer disease. Thirty-nine analyses were performed for the comparator studies; NNTB ranged from -61.8 to 330.8 (median = 4.6). The median value of NNTB values was lower for ERT studies than for comparator studies, suggesting that ERT therapies for rare diseases were more effective than existing recommended drugs for comparable diseases. Caution should be applied to the interpretation of these results because the analyses were limited by risk of bias, study size and the lack of identical outcomes. Comparing the effectiveness of orphan drugs with non-orphan drugs using NNTB may provide additional clinical effectiveness information for reimbursement decision making.

KEYWORDS

enzyme replacement therapy, lysosomal storage disorders, rare diseases, effectiveness.

CORRESPONDING AUTHOR:

Shona Lang
Kleijnen Systematic Reviews Ltd., Unit 6, Escrick Business Park, Riccall Road, Escrick, York, YO19 6FD
Tel. +44(0)1904 727992
e-mail: shona@systematic-reviews.com

Submitted: 20 January 2014 / Accepted: 31 March 2014

Copyright © 2014 Shona Lang et al; licensee Istituto Superiore di Sanità. This is an open access article licensed under the Creative Commons license Attribution-Noncommercial 3.0 Unported, which permits to copy and redistribute the material in any medium or format and remix, transform, and build upon the material for non commercial use and provided that the original work is properly cited. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc/3.0/

The Rare Diseases and Orphan Drugs Journal has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under Grant Agreement n° 305690: RARE-Bestpractices project: www.rarebestpractices.eu. Sole responsibility lies with the authors and the European Commission is not responsible for any use that may be made of the information contained therein.
WWW.RAREBESTPRACTICES.EU

INTRODUCTION

Enzyme replacement therapy (ERT) is a medical treatment to reintroduce an enzyme into a patient deficient for that enzyme. ERT is usually performed by intravenous infusion. Lysosomal storage disorders are a group of genetically inherited metabolic diseases that result in the deficiency of lysosomal enzymes. Lysosomes are sub cellular organelles which play a central role in the normal recycling of the cell and contain enzymes to degrade cellular macromolecules[1]. The absence of these enzymes leads to the accumulation of unwanted molecules within a cell which then interfere with the normal function of that cell and ultimately a whole tissue. Although they share a basic pathological mechanism, LSDs are clinically remarkably heterogeneous and can affect completely different tissues. Lysosomal storage disorders are broadly divided into groups based on the nature of the stored substrate: mucopolysaccharidoses (MPS), lipidoses, glycogenoses, and oligosaccharidoses[2]. Lysosomal diseases are rare, each being classified as an orphan disease (prevalence less than five in 10,000 in Europe), although together their incidence is about one in 6,500 to 7,500, similar to that of Cystic Fibrosis (CF), one of the commonest genetic diseases [3].

ERT is available for several lysosomal storage disorders including Fabry Disease, Hunter Syndrome (MPS II) and Gaucher Disease Type 1. Fabry Disease is caused by a deficiency of the lysosomal enzyme alpha-galactosidase A (AGAL). Clinically, Fabry Disease is characterised by major renal, cardiac and cerebrovascular complications caused by a build up of cellular globotriaosylceramide [4]. The life expectancy for patients is 50 for men and 70 for women whilst the incidence is 1 in 117,000 live births for males [5].

MPS II or Hunter Syndrome is caused by a deficiency in iduronate-2-sulphatase leading to an accumulation of dermatan sulphate and heparan sulphate in nearly all cell types, tissues and organs. Clinically Hunter Syndrome has multi-organ and multi-system involvement with a variable age of onset and rate of progression. Evidence of more severe forms in some patients emerges between ages 2-4 years, with severe neurologic involvement and death usually occurs in the first or second decade of life, usually due to obstructive airway disease and/or cardiac failure. Patients with milder forms have a later onset, minimal neurological dysfunction and often survive into adulthood. The incidence of MPS II is estimated between 1 in 162,000 and 1 in 170,000 male live births [5,6].

Gaucher disease is caused by a deficiency in the enzyme glucocerebrosidase leading to an accumulation of glucocerebroside in cells of the immune system. Clinically Gaucher Disease is characterised by: visceral problems (hepatomegaly, splenomegaly, anaemia and thrombocytopenia) causing fatigue, discomfort, infections, bleeding and bruising; bone problems (pain and necrosis); lung disease, impaired growth and delayed puberty. Type I can present at any age. The incidence of Gaucher disease is 1 in 40,000 to 60,000 live births and has a life expectancy of 68 years [7].

Currently there is FDA marketing approval in the United States (http://www.fda.gov/default.htm) and EMA approval in Europe (http://www.ema.europa.eu/ema/) for a combined total of eight ERTs for different lysosomal storage disorders. These include: agalsidase alpha or agalsidase beta for Fabry Disease; Idursulfase for Hunter Syndrome (MPS II); velaglucerase alpha and imiglucerase for mild to moderate Type I Gaucher Disease.

Due to their low prevalence, rare diseases have traditionally been neglected by industry and by the scientific, medical and political communities. Therefore, in the United States and in the European Union, incentives have been put in place to stimulate the development of 'orphan drugs', that is, drugs developed to treat these orphan or rare diseases, by compensating industry for the risks and lower potential return on investment as a consequence of the inherently low number of patients. In addition networking has been established to improve databases and clinical trial expertise [8].

Orphan diseases often have lengthy time courses, irreversible progression and few, clinically variable patients. Therefore the use of specific clinical measures as endpoints for ERTs can make drug development difficult or challenging for practical and ethical reasons. However by law, all drugs must undergo clinical trial testing to demonstrate safety and substantial efficacy before FDA or EMA approval. This process usually requires a phase 3, double-blind, placebo controlled trial, widely regarded as the "gold standard". Consequently it has been suggested that orphan drug approval should be allowed on the use of surrogate endpoints "reasonably likely to predict clinical benefit" - a surrogate endpoint being a measure, such as a blood test or urine marker, believed to be indicative of a disease state and treatment effect, but not demonstrative of a direct health gain to the patient [9]. Therefore endpoints used to study effectiveness in rare diseases must be carefully considered for their ability to represent the disease and for their use in comparative studies.

Reimbursement of a new drug by a health care system occurs after an assessment of value for money which compares the cost and clinical effectiveness of the new technology to that of an existing technology. For lysosomal storage disorders and other rare diseases there are often no effective standard treatments to use as comparators. Costs per quality adjusted life year are always likely to be high. One approach to put clinical effectiveness further into perspective could be to compare numbers needed to treat for ERT to those of drugs which do get reimbursed for comparable diseases with greater numbers of patients. This could demonstrate that an ERT has relatively large effects for patients who have no alternative treatment available. However, a challenge is then to define comparable diseases, in terms of impact on quality of life, chronicity and course of disease (including reversibility or irreversibility).

There are a lack of randomised controlled trials and a lack of agreed methodology for the study of ERT effectiveness as required by reimbursement agencies. This paper aims to highlight some of the issues surrounding the study

of ERT effectiveness and to provide some suggestions for methodology to develop this area. We set out to establish the effectiveness of ERT for Fabry Disease, Hunter Syndrome (MPS II) and Gaucher Disease Type 1 in comparison to drugs which have received a recommendation reimbursement by NICE (UK) or IQWiG/G-BA (Germany) in 2012 in comparable diseases with greater prevalence. To compare effectiveness between diseases we compared the number needed to treat for an additional patient to have a beneficial outcome for the primary outcome (or matching outcomes) of each trial (NNTB).

METHODS

Inclusion criteria for ERT studies

Randomised controlled trials or meta analyses of patients with Fabry Disease, Gaucher Disease type I or Hunter Syndrome (MPS II) who were treated with ERT in comparison to placebo were included in the systematic review.

Literature searches for ERT studies

Focused searches were undertaken to identify relevant trials and systematic reviews of the effectiveness of five ERTs for the treatment of Hunter Syndrome (Idursulfase), Gaucher Disease type I (imiglucerase, velaglucerase alpha) and Fabry Disease (agalsidase alpha and beta) regardless of language. The search strategies (keywords) were developed specifically for each database (example shown in Additional file 1). Searches took into account generic and other product names for drugs including variations in different countries. Only studies conducted in humans were sought. Specific search filters for randomised controlled trials were used to retrieve studies of clinical effectiveness. To increase sensitivity, the literature search was broadened to include other rare diseases (Hurlers Syndrome, Pompe disease and MPS VI) which also receive enzyme replacement therapy (in case of misclassification or reporting of the included diseases within publications of the additional diseases).

The following databases were searched from inception up to August 2012: MEDLINE, MEDLINE In-Process Citations and Daily Update, EMBASE, Cochrane Database of Systematic Reviews, Cochrane Central Register of Controlled Trials, Database of Abstracts of Reviews of Effects, Health Technology Assessment Database, ClinicalTrials.gov, metaRegister of Controlled Trials (mRCT) and the WHO International Clinical Trials Registry Platform (ICTRP).

The main EMBASE strategy was independently peer reviewed by a second Information Specialist, using the PRESS-EBC checklist [10]. Identified references were downloaded in Endnote X4 software for further assessment and handling. References in retrieved articles were checked for additional studies and the final list of included papers was also checked on PubMed for retractions and errata [11-13].

Methods of study selection, quality assessment and data extraction

Two reviewers independently inspected the title and abstract of each reference identified by the search to determine the potential relevance and eligibility of each article according to the criteria specified above. For potentially relevant articles, or in cases of disagreement, the full article was obtained, independently inspected, and inclusion criteria applied. Any disagreements were resolved through discussion and consensus.

Data were extracted by one reviewer and checked by a second reviewer. The following main information was extracted from studies: author, year, country, diagnosis, predefined inclusion and/or exclusion criteria, duration of follow up, description of the participants included in the study, study aim and conclusions, treatment and dose, number of participants recruited/included/withdrawn in the study. All outcomes were extracted. Dichotomous data were extracted as the number of individuals with the outcome of interest and the total numbers of individuals in the intervention and control group. Continuous data were extracted as the mean and standard deviation (SD) for the intervention and control group.

Assessment of methodological quality was based on the Cochrane Collaboration checklist [14] and AMSTAR [15].

Selection and extraction of comparator drugs

The technology appraisal zone of the NICE and G-BA websites were searched for drugs that NICE or G-BA has recommended for reimbursement between 2010 and 2012: http://guidance.nice.org.uk/TA/Published; http://www.g-ba.de/. From this list, the following inclusion criteria were applied: the diseases had a prevalence that was greater than that of an orphan disease (>5 persons in 10,000 of the population), drug was compared to placebo; patients were likely to receive lifelong treatment; and were likely to have a shortened life expectancy due to the disease.

All documents on the websites were examined to provide as much relevant information as possible (primarily the technology appraisal and the evidence review). All outcomes were extracted. If relevant or comparable outcomes (to the ERT) were identified then the original papers were also accessed if data were not available in the original appraisal. Quality assessments were extracted from the evidence review group data.

Data analysis

All reported outcomes were extracted from each study. The analysis aimed to identify efficacy outcomes which were reported by both ERT and non-ERT comparator studies. In the absence of identical outcomes one outcome was selected for analysis as follows: 1) for ERT studies, we first selected any outcomes which were comparable between ERT studies and then we chose the primary outcome or the outcome which was reported first in the study abstract. 2) For comparator studies we first selected the primary outcome according to the appraisal document but if this was a continuous outcome (or did not report standard devia-

tions or standard errors) we selected the first reported secondary, dichotomous outcome. For a given outcome all follow-up times were analysed and data from all trials providing evidence to the guidance were used.

The number needed to treat for an additional patient to have a beneficial outcome (NNTB) with 95% confidence intervals (CI) for treatment compared with standard care/ placebo was calculated for each relevant outcome (at all treatment doses) for each study and to allow exploratory comparisons between the ERT and comparator studies.

For dichotomous outcomes NNTB was calculated as 1/absolute risk difference for treatment versus placebo. Continuous outcomes needed to be converted into a dichotomous format to allow NNTB calculations, so the following assumptions were made:

We assumed that the continuous variable was normally distributed within the sample. We then estimated the probability of this variable being below a particular threshold from the cumulative normal distribution (P, the probability that the value is below a particular threshold, parameterised by the mean and SD). Therefore, if we then assume that the threshold determines the treatment response then we have either P (if treatment response is below the threshold) or 1-P (if the treatment response is above the threshold) which correspond to the probability of a response or non-response.

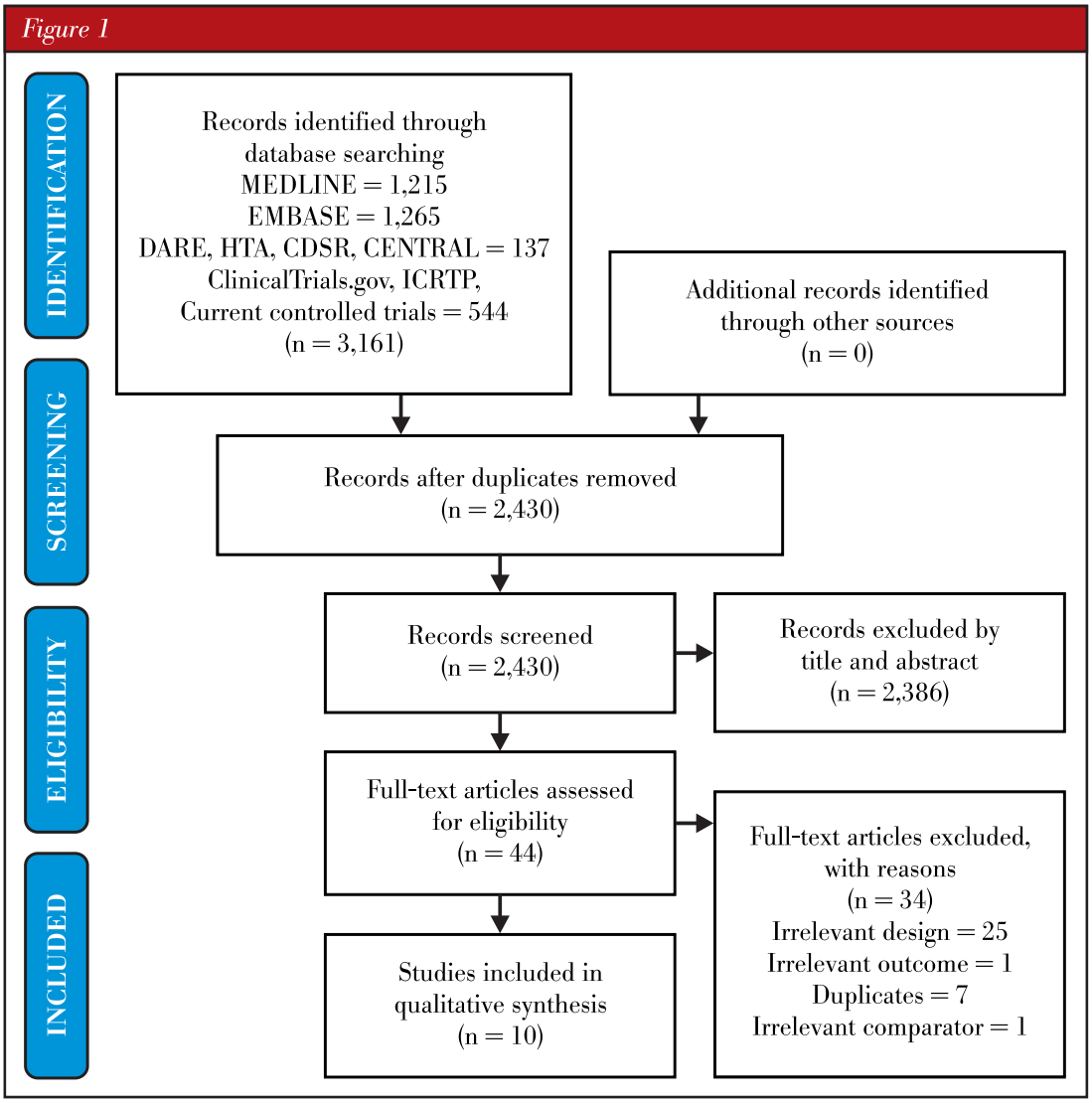
We then had to judge what the threshold should be. We assumed that any change from baseline (in a positive direction; i.e. consistent with slowing of disease progression) would be a treatment response. Therefore, where we had data for the change from baseline to follow-up the threshold was assumed to be zero (indicating no change) and where we had only data on follow-up, we used the mean baseline value across treatment and placebo arms.

Note that for NNTB the value is not always contained within the confidence intervals; this is due to calculating NNTB as 1/absolute risk difference, therefore $1/0 = \infty$ [16]. A narrative synthesis method was used to summarise the data and a detailed commentary of the major methodological problems or biases that affected the studies was also included. Results for NNTB were presented in figures and tables grouped by disorder/condition and ERT versus non-ERT treatments.

RESULTS AND DISCUSSION

For the ERT studies, our searches of the databases yielded 2,430 articles (after de-duplication). 2,386 were excluded after screening the titles and abstracts and 44 full text articles were screened. Of these 44 records, 10 publications of randomised controlled trials were identified *(Figure 1)*.

Figure 1



Flow of the information for the enzyme replacement therapy (for Hunter Syndrome, Gaucher Disease type I and Fabry Disease) literature search through the different phases of the systematic review.

Key study and baseline characteristics were extracted, and are summarised in *Table 1*.

Table 1: Characteristics of randomised controlled trials of enzyme replacement therapies for fabry disease and hunter syndrome (Continues)

First author & Publication year	Population	ERT	Dose	Follow-up (m)	Outcomes	Basic characteristics
Banikazemi 2007[19]	Fabry	β galactosidase vs. placebo	1 mg/kg EOW	35	Composite outcome[p]* (cardiac and cerebrovascular outcomes), SAE, AE, TAE, death	% Male: 88 Mean Age (range): 45 (NR) N=82
Bierer 2006[20]	Fabry	α galactosidase vs. placebo	1 mg/kg EOW	Minimum 18	Max oxygen uptake at peak exercise[p]*, heart rate reserve, estimated stroke volume, respiratory rate at peak exercise, ventilatory reserve	% Male: 83 Mean Age (range): 32 (20-47) N=6
Eng 2001[21]	Fabry	β galactosidase vs. placebo	1 mg/kg EOW	5	Renal capillary endothelial clearance of Gb3[p]*, pain (VAS score).	% Male: 97 Mean Age (range): 30 (16-61) N=58
Schiffman 2001[22]	Fabry	α galactosidase vs. placebo	0.2 mg/kg/ EOW	2,4,6	BPI Pain severity[p], BPI pain related QOL (McGill), days without medication, normal glomeruli, creatinine clearance, insulin clearance, plasma Gb3*, body weight	% Male: 100 Mean Age (range): 34 (NR) N=26
Schiffman 2006[23]	Fabry	α galactosidase vs. placebo	0.2 mg/kg/ EOW	6	Intraepidermal nerve fibre density in thigh[p]*	% Male: 100 Mean Age (range): 34 (NR) N=26
Moore 2002 [24]	Fabry	α galactosidase vs. placebo	Dose NR/ EOW	6	Regional cerebral flow[p]*, global cerebral flow	% Male: 100 Mean Age (range): 33.7 (19-48) N=26
Hajioff 2003[25]	Fabry	α galactosidase vs. placebo	0.2 mg/kg/ EOW	6	Change in high frequency hearing loss[p+]	% Male: 100 Mean Age (range): NR (16-56) N=15
Hughes 2008[26]	Fabry	α galactosidase vs. placebo	0.2 mg/kg/ EOW	6	Myocardial Gb3 levels[p], left ventricular mass, left ventricular ejection fraction, plasma Gb3 levels*, SAE	% Male: 100 Mean Age (range): 37 (23-51) N=15
Muenzer 2007 [17]	Hunter	Idursulfase Vs placebo	0.15, 0.5, 1.5 mg/kg/EOW	12	Urinary GAG excretion[p]*, 6MWT distance	% Male: 100 Mean Age (range): 14(6-20) N=12

Characteristics of randomised controlled trials of enzyme replacement therapies for fabry disease and hunter syndrome (Continued)						
Muenzer 2006[18]	Hunter	Idursulfase Vs placebo	0.5 mg/kg/wk Or 0.5mg/kg/ EOW	6	Composite outcomes[p] (Urinary GAG excretion, 6 minute walk test distance), Urinary GAG excretion*, 6MWT distance, predicted force vital capacity, absolute force vital capacity, liver volume, spleen volume, AE, SAE, deaths	% Male: NR Mean Age (range): 14 (5-31) N=96

*[p]primary outcome or first reported outcome; *outcome used to calculate NNTB; [¤] outcome could not be converted to NNTB; EOW = every other week; wk=weekly; m= months; ERT = enzyme replacement therapy; N= number of patients; NR = not reported; SAE = serious adverse events; AE= adverse events; TAE = treatment related adverse events; QOL = quality of life; GAG= glycomsaminoglycan; Gb3= globotriaosylceramide; 6MWT = 6 minute walk test; VAS = visual analogue scale; BPI = brief pain inventory.*

Two studies [17, 18] investigated patients with Hunter Syndrome, eight [19-26] investigated patients with Fabry Disease and no studies were found for Gaucher Disease type I. The studies for Hunter Syndrome compared idursulfase (0.15, 0.5, 1.5 mg/kg) with placebo; studies of Fabry disease compared alpha galactosidase (0.2, 1mg/kg) or beta agalsidase (1mg/kg) with placebo. Follow-up times ranged from 2 to 35 months. The outcomes reported varied considerably between the different diseases and within the same disease. Two Fabry disease studies [26, 22] reported plasma Gb3 levels and both Hunter disease studies [17, 18] reported Urinary GAG excretion and 6 minute walking distance, but no other outcomes could be matched. The majority of patients were male. In Fabry disease the age of the patients ranged from 16 to 61 years old and in Hunter Syndrome the patients' ages ranged from 5 to 31 years.

For the comparator studies, 47 appraisals were identified from NICE and G-BA websites and from this list seven met the inclusion criteria (two were identified from the G-BA website but these were already identified from NICE and were not included). For each of the seven appraisals we included the guidance and the evidence review group report. For two of the appraisals (TA203, TA225) we had to identify the original trial reports to extract all data and so four additional papers were included [28-31]. A total of 18 reports (seven guidance reports, seven ERG reports, four original trial reports) were identified and are listed in *(Table 2)*.

Table 2: Characteristics of comparator studies of reimbursed drugs (Continues)						
Technology Appraisal	Data sources	Population	Drugs recommended by GBA/NICE	Dose (mg)	Follow-up (weeks)	Outcomes
TA254	NICE guidance[40] Appraisal [33]	Multiple sclerosis (relapsing-remitting)	Fingolimod	0.5	13, 26	Rate of relapse[p], absence of disability progression*, time to relapse, MRI outcomes SAE, death
TA198	NICE guidance [41] Appraisal [32]	Rheumatoid arthritis	Tocilizumab (plus methotrexate)	8	24, 52	ACR20[p]* ACR50, ACR70, HAQ, DAS-28, SAE
TA203	NICE guidance[42] Appraisal LEAD 4[28]	Type 2 diabetes mellitus	Liraglutide (in combination with metformin and a sulphonylurea, or metformin and a thiazolidinedione)	1.2	26	Change in HbA1c[p], patients reaching HbA1c < 7.0%*, haemoglobin levels, blood glucose, fasting glucose, nausea.

Characteristics of comparator studies of reimbursed drugs (Continued)						
TA225	NICE guidance [43] Appraisal [44] Kay 2008[29] GO-FORWARD[30] GO-AFTER[31]	Rheumatoid arthritis (after the failure of conventional drugs)	Golimumab (plus methotrexate)	50, 100	14-16, 24	ACR20[P]*, ACR50, ACR70, ACR90, HAQ, DAS-28, Injection site reactions, infections, withdrawals, SAE, death
TA223	NICE guidance [45] Appraisal [46]	Peripheral arterial disease	Naftidrofuryl Oxalate	600	24	Maximum walking distance[P], pain-free walking distance*, ankle-brachial pressure index, vascular events, requirement of hospitalisation, HRQOL, TAE, SAE, death
TA195	NICE guidance [47] Appraisal [48]	Rheumatoid arthritis (severe active)	Rituximab	2,000	12, 24	ACR20[P]*, ACR50, ACR70, disease activity (DAS), physical function (HAQ), infections, HRQOL (SF36), SAE, death, withdrawals.
			Abatacept	10 mg/kg		
TA217	NICE guidance [49] Appraisal [27]	Mild to moderate Alzheimer's disease	Donepezil	10	12, 12-16, 21-28, 24	SIB*[P], cognition (ADAS-cog)*[P], MMSE, functional (ADL, DAD, GAS-VR), behaviour and mood, clinical dementia ratings, mortality, ability to remain independent, likelihood of admission to residential/nursing care, HRQOL of patients and carers, AEs of treatment, SAE, death
			Galantamine	≤24		
			Rivastigmine	≥12		
		Severe Alzheimer's disease or moderate disease (if intolerant or contraindicated to ACE inhibitors)	Memantine	5-20		

*[P]primary outcome or first reported outcome according to appraisal; *outcome used to calculate NNTB; [2] outcome could not be converted to NNTB; EOW = every other week; wk=weekly; m= months; ERT = enzyme replacement therapy; N= number of patients; NR = not reported; SAE = serious adverse events; AE= adverse events; TAE = treatment related adverse events; QOL = quality of life; GAG= glycosaminoglycan; GB3= globotriaosylceramide; ACR(20)=American College of Rheumatology score (patients have ≥ 20% fewer tender joints and ≥ 20% fewer swollen joints); HbA1c = glycated haemoglobin; HAQ= Health Assessment Questionnaire; DAS = disease activity score; MMSE = minimal mental state examination score; SIB= severe impairment battery score; ADAS = Alzheimer›s Disease Assessment Scale score.*

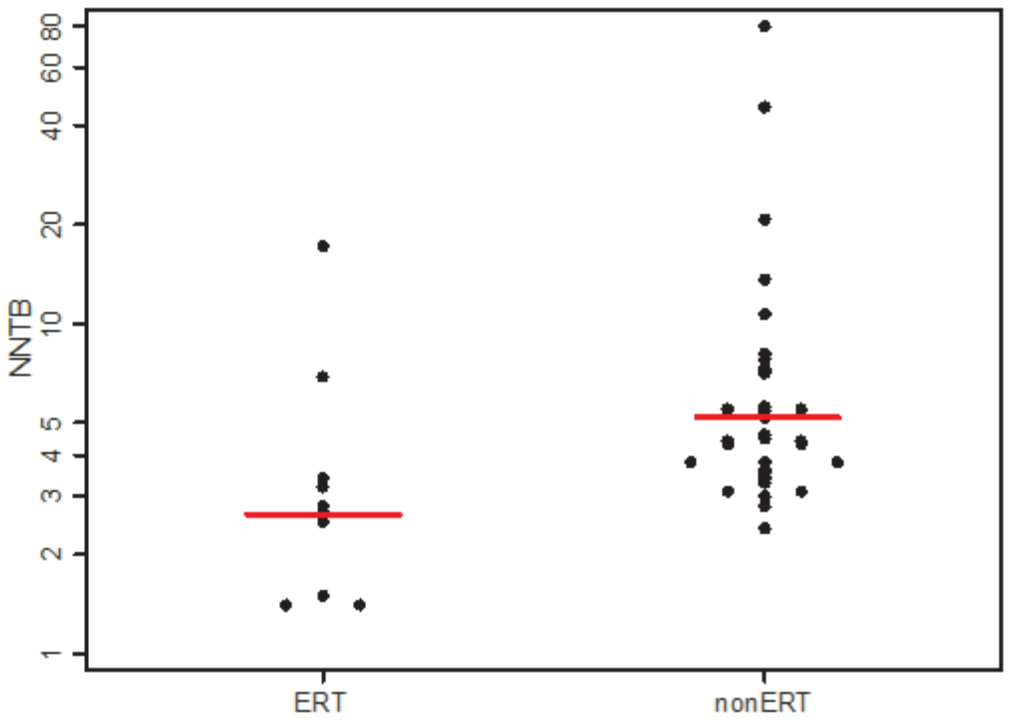
Appraisals from G-BA did not yield any additional comparators. Diseases that were judged to be comparable to orphan disease were; multiple sclerosis, rheumatoid arthritis, type 2 diabetes mellitus, peripheral arterial disease and Alzheimer disease. In total, eleven comparator drugs were identified (rituximab, abatacept, golimumab, fingolimod, liraglutide, naftidrofuryl, tocilizumab, donepezil, galantamine, rivastigmine, memantine). Follow-up times varied from 12 to 52 weeks and outcomes varied considerably between diseases.

The quality of the ERT studies and non ERT/comparator studies is summarised in Appendix 2. For the ERT studies,four studies [22-25] included at least one source of high risk of bias, due to incomplete outcome data reporting. The remaining studies had no source of high risk bias; however none of the studies clearly reported all areas of the quality assessment making it difficult to accurately judge the overall risk of bias for any study. For the non-ERT/comparator studies, the quality (as presented in the technology assessment) for each individual trial was assessed. Only one of the reviews (TA198 [32]) did not perform a quality assessment. One study clearly reported all areas of the quality assessment and had a low risk of bias [28]. Of the remaining five

reviews, only two [27, 33] included at least one source of high risk of bias. A comprehensive literature search was performed to reduce the risk of publication bias. Formal investigation of publication bias was not possible because there were too few studies of variable size [34]. However, this bias cannot be ruled out, especially the effects of negative publication, whereby results in the field are not published due to lack of data, interest or positive results. The field of rare diseases is particularly susceptible to a lack of data.

Forty one reported outcomes were identified from the ERT studies and sixty one for the comparator studies. None of the ERT outcomes for effectiveness matched those for the comparator studies, although both presented safety data (adverse events and deaths). Within the ERT studies we identified very few studies reporting the same outcome; two studies reported urinary GAG excretion and 6 minute walking distance [17, 18] and two studies reported plasma Gb3 levels [26, 22]. There were insufficient study numbers to perform meta analysis. To establish the effectiveness of the ERT therapy, for each study we selected one outcome which was either comparable to another study (urinary GAG, plasma GB3) or a primary outcome (urinary GAG, cardiac and cerebrovascular outcome, oxygen uptake, microvascular deposits, cerebral flow), as indicated in *Table 1*. Only two of the reported effectiveness outcomes were dichotomous (cardiac and cerebrovascular outcome, microvascular deposits), no other dichotomous outcomes were identified. Two studies were rejected for analysis; one study did not report mean and standard deviation [25] and one study was rejected because the only outcome (nerve fibre density) was not considered a good assessment for Fabry disease [23]. For all other outcomes an anlaysis was performed for each reported dose and follow-up time. Eleven NNTB analyses (plus 95% confidence intervals) were performed and are presented in Additional File 3 and summarised in *(Figure 2)*.

Figure 2



Comparison of NNTB values derived from enzyme replacement therapy (ERT) studies and comparator disease studies (non ERT). Each circle indicates an individual NNTB (number needed to treat in order to benefit one person). The red line indicates the median value. Eleven analyses were performed for ERT studies and thirty nine were performed for non-ERT studies.

To establish the effectiveness of the comparator therapies, for each study we selected one outcome which was either comparable to another study (ACR20) or a primary outcome (ACR20, ADAS, SIB) or a secondary dichotomous outcome if the primary were continuous (absence of disability progression, HbA1c <7%) or the primary did not report standard deviations (pain free walking distance), as indicated in *Table 2*. For each outcome an analysis was performed on each approved dose and for each follow-up time. Thirty-nine analyses (Additional File 3) were performed and are summarised in Figure 2. For the purpose of summarising the data in Figure 2 we removed any outliers; NNTB >100 or negative values (3 values from the non-ERT comparators).

For all ERT studies NNTB calculated from the mean absolute risk difference ranged from 1.4 to 17.2, the median value was 2.7 (mean = 4.15, standard deviation = 4.59). Analysis of the individual drugs found that the use of agalsidase beta produced NNTB values of 1.5 and 6.9, agalsidase alpha produced values of 2.7 to 17.2 (median = 3.3) and idursulfase produced values of 1.4 to 2.8 (median = 2.5).

For comparator studies NNTB calculated from the mean absolute risk difference ranged from -61.8 to 330.8, the median value was 4.6 (mean = 15.99, SD = 55.05). For the comparator analyses there were two negative values for NNTB (see Additional File 3) this indicated that the treatment had a negative effect (had no benefit) compared to the control.

Conversion of continuous outcomes to NNTB values is controversial [35, 36] therefore we analysed only the dichotomous

outcomes (noted in Additional File 3). However, only two analyses used dichotomous outcomes for the ERT studies, the median value was 4.2. Twenty two analyses used dichotomous outcomes in the comparator studies, and the median value was 4.05 (range; 2.4 to 79.8). No conclusions can be drawn due to the low number of ERT analyses.

To try to reduce the heterogeneity of the results we examined studies with follow-up times of 4 to 6 months. This produced similar results to those overall; NNTB for ERT studies (n=7) ranged from 1.5 to 17.2 with a median value of 2.8 and for comparators (n=26) NNTB ranged from -61.8 to 79.8, median was 4.4. There was insufficient data to repeat this analysis for 12 months.

Overall, the median value of NNTB values was lower for ERT studies than for comparator studies, suggesting that ERT therapies for rare diseases are more effective than non-ERT drugs for comparable but more prevalent diseases (however the result is not significant). The use of NNTB may provide additional effectiveness evidence to support the reimbursement of ERT therapy by indicating relatively greater clinical effectiveness despite higher costs.

The results were limited by the quality of the original studies and the heterogeneity of the outcomes. More than half of the ERT studies were considered of unknown quality due to a lack of clear methodological reporting. Additionally the ERT studies were based on those for rare diseases which have small patient populations and are therefore inherently biased due to their size and difficulties associated with the reporting of rare diseases [37, 38]. Studies with small numbers of patients have low statistical power and reduce the reliability of NNTB calculations. To convert continuous data into NNTB values the data are assumed to follow a normal distribution, however this assumption cannot be validated with low patient numbers. Comparisons of NNTB data derived from ERTs and comparator studies were limited by the lack of identical outcomes. This was to be expected since the non-ERT studies were from different patient populations, whose disease states were monitored with different outcomes. It is also true that NNTB and NNTH are limited as outcome measures insofar as some reimbursement agencies e.g. NICE often use QALYs. Therefore, we would not suggest that NNTH and NNTB should replace QALYs, but instead supplement the ways of comparing technologies. Indeed the latest version of the guide to the process for the evaluation of Highly Specialised Technologies (essentially orphan drugs) [39] does not mention QALYs, but does suggest a potential role for NNTB in the form of 'Overall magnitude of health benefits', which is one of the criteria recommended for consideration by the committee.

Future work should investigate the selection of comparators based on the identification of identical/similar outcomes rather than on the basis of disease comparability (similar quality of life and chronicity) as was done here, or a combination of both. Overcoming the limits of rare disease research requires the collaboration of researchers and clinicians of these very specific diseases. Future trials require greater consistency in the reporting of the research methods; co-ordinated use of the best outcomes for each disease and further co-ordination to create trials that are as large as possible, to ensure that future results will be more widely useful. The design of future clinical trials for orphan diseases is currently being assessed by the 'European Reference Networks for Rare Diseases' [8] and hopefully they will tackle some of the limitations highlighted here.

Despite the limitations of the evidence provided here the use of NNTB has allowed us to compare the effectiveness across multiple outcomes and diseases to try and achieve the aims of our research. It can be argued that we have provided useful clinical effectiveness information for the reimbursement procedure.

CONCLUSIONS

The median value of NNTB values was lower for ERT studies than for comparator studies (2.7 compared to 4.6), suggesting that ERT is more effective for rare diseases than existing recommended drugs for more prevalent but comparable diseases. Caution should be applied to the interpretation of these results because the analyses were limited by risk of bias due to the lack of reporting standards, study size, heterogeneous outcomes and the conversion of continuous outcomes to estimate NNTB. However, it seems feasible to compare the effectiveness of orphan drugs with non-orphan drugs using NNTB. Such analyses will provide improved clinical effectiveness data to support the reimbursement decision making process.

List of Abbreviations: NNTB = number needed to treat for an additional beneficial outcome to one patient; EOW = every other week; wk=weekly; m= months; ERT = enzyme replacement therapy; N= number of patients; NR = not reported; SAE = serious adverse events; AE= adverse events; TAE = treatment related adverse events; QOL = quality of life; GAG= glycomsaminoglycan; GB3= globotriaosylceramide; SD = standard deviation; 6MWT = 6 minute walk test; VAS = visual analogue scale; BPI = brief pain inventory; ACR(20)=American College of Rheumatology score (patients have ≥ 20% fewer tender joints and ≥ 20% fewer swollen joints); HbA1c = glycated haemoglobin; HAQ= Health Assessment Questionnaire; DAS = disease activity score; MMSE = minimal mental state examination score; SIB= severe impairment battery score; ADAS = Alzheimer's Disease Assessment Scale score.

Authors' contributions: SL participated in the systematic review, performed the statistical analysis and drafted the manuscript. NA participated in the design of the study, participated in the systematic review and performed statistical analyses. JK participated in the design of the study and participated in the systematic review. CN performed the literature search. All authors read and approved the final manuscript.

Acknowledgements: All the authors and the manuscript were funded by Charles River Associates (who were funded by Shire plc). The funders did not contribute to the design, analysis, interpretation or in the submission of the manuscript. The authors would like to thank Kate Misso for her assistance in drafting the manuscript with regards to the literature searching.

REFERENCES

1. Platt FM, Boland B, van der Spoel AC: **Lysosomal storage disorders: The cellular impact of lysosomal dysfunction.** *The Journal of Cell Biology* 2012,199:723-734.
2. Lachmann RH: **Enzyme replacement therapy for lysosomal storage diseases.** *Curr Opin Pediatr* 2011,23:588-593 10.1097/MOP.0b013e32834c20d9.
3. National Institutes of Health Clinical Center (CC), National Institute of Neurological Disorders and Stroke (NINDS) (2009): **OGT 918-006: a phase I/I randomized, controlled study of OGT 918 in patients with neuronopathic Gaucher Disease.** *NIH Clinicaltrials.gov*. http://ClinicalTrials.gov/show/NCT00041535
4. El Dib RP, Pastores GM **Enzyme replacement therapy for Anderson-Fabry Disease**. *Cochrane Database of Systematic Reviews 2010, Issue 5. Art. No: CD006663. DOI: 10.1002/14651858.CD006663.pub2*. http://www.mrw.interscience.wiley.com/cochrane/clsysrev/articles/CD006663/frame.html
5. Meikle PJ HJ, Clague AE, Carey WF: **Prevalence of lysosomal storage disorders.** *JAMA* 1999,281:249-254.
6. Martin R, Beck M, Eng C, Giugliani R, Harmatz P, Muñoz V, *et al.*: **Recognition and Diagnosis of Mucopolysaccharidosis II (Hunter Syndrome).** *Pediatrics* 2008,121:e377-e386.
7. Connock M, Burls A, Frew E, Fry-Smith A, Juarez-Garcia A, McCabe C, *et al.*: **The clinical effectiveness and cost-effectiveness of enzyme replacement therapy for Gaucher's Disease: a systematic review.** *Health Technol Assess* 2006,10:1-152.
8. European Commission: **Public health [Internet]**[http://ec.europa.eu/dgs/health_consumer/about_us/contact_en.htm]
9. Miyamoto BE, Kakkis ED: **The potential investment impact of improved access to accelerated approval on the development of treatments for low prevalence rare diseases.** *Orphanet J Rare Dis* 2011,6:49.
10. Sampson M, McGowan J, Cogo E, Grimshaw J, Moher D, Lefebvre C: **An evidence-based practice guideline for the peer review of electronic search strategies.** *J Clin Epidemiol* 2009,62:944-52.
11. Royle P, Waugh N: **Should systematic reviews include searches for published errata?** *Health Info Libr J* 2004,21:14-20.
12. Wright K, McDaid C: **Is the retraction of journal articles in electronic journals and databases consistent and timely? A case study [Poster]**. Paper presented at *Cochrane Colloquium; 19-22 October 2011; Madrid*. In *Cochrane Collaboration*, 2011.
13. Wright K, McDaid C: **Reporting of article retractions in bibliographic databases and online journals.** *J Med Libr Assoc* 2011,99:164-7.
14. Higgins JPT, Green S, eds: *Cochrane handbook for systematic reviews of interventions [Internet]*. Version 5.1.0 [updated March 2011]: The Cochrane Collaboration, 2011 [accessed 23.03.11]. Available from: http://www.cochrane-handbook.org/
15. Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, *et al.*: **Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews.** *BMC Med Res Methodol* 2007,7:10.
16. Altman DG: **Confidence intervals for the number needed to treat.** *BMJ* 1998,317:1309-1312.
17. Muenzer J, Gucsavas-Calikoglu M, McCandless SE, Schuetz TJ, Kimura A: **A phase I/II clinical trial of enzyme replacement therapy in mucopolysaccharidosis II (Hunter Syndrome).** *Mol Genet Metab* 2007,90:329-37.
18. Muenzer J, Wraith JE, Beck M, Giugliani R, Harmatz P, Eng CM, *et al.*: **A phase II/III clinical study of enzyme replacement therapy with idursulfase in mucopolysaccharidosis II (Hunter Syndrome).** *Genet Med* 2006,8:465-73.
19. Banikazemi M, Bultas J, Waldek S, Wilcox WR, Whitley CB, McDonald M, *et al.*: **Agalsidase-beta therapy for advanced Fabry Disease: a randomized trial.** *Ann Intern Med* 2007,146:77-86.
20. Bierer G, Balfe D, Wilcox WR, Mosenifar Z: **Improvement in serial cardiopulmonary exercise testing following enzyme replacement therapy in Fabry Disease.** *J Inherit Metab Dis* 2006,29:572-9.
21. Eng CM, Guffon N, Wilcox WR, Germain DP, Lee P, Waldek S, *et al.*: **Safety and efficacy of recombinant human alpha-galactosidase a replacement therapy in Fabry's Disease.** *N Engl J Med* 2001,345:9-16.
22. Schiffmann R, Kopp JB, Austin IHA, Sabnis S, Moore DF, Weibel T, *et al.*: **Enzyme replacement therapy in Fabry Disease a randomized controlled trial.** *JAMA* 2001,285:2743-9.
23. Schiffmann R, Hauer P, Freeman B, Ries M, Scott LJC, Polydefkis M, *et al.*: **Enzyme replacement therapy and intraepidermal innervation density in Fabry Disease.** *Muscle Nerve* 2006,34:53-6.
24. Moore DF, Altarescu G, Herscovitch P, Schiffmann R: **Enzyme replacement reverses abnormal cerebrovascular responses in Fabry Disease.** *BMC Neurology* 2002,2.
25. Hajioff D, Goodwin S, Quiney R, Zuckerman J, MacDermot KD, Mehta A: **Hearing improvement in patients with Fabry Disease treated with agalsidase alfa.** *Acta Paediatr Suppl* 2003,92:28-30.
26. Hughes DA, Elliott PM, Shah J, Zuckerman J, Coghlan G, Brookes J, *et al.*: **Effects of enzyme replacement therapy on the cardiomyopathy of Anderson-Fabry Disease: a randomised, double-blind, placebo-controlled clinical trial of agalsidase alfa.** *Heart* 2008,94:153-8.
27. Bond M, Rogers G, Peters J, Anderson R, Hoyle M, Miners A, *et al.*: **The effectiveness and cost-effectiveness of donepezil, galantamine, rivastigmine and memantine for the treatment of Alzheimer's Disease (review of TA111): a systematic review and economic model** *[Internet]*. Exeter: Peninsula Technology Assessment Group (PenTAG), University of Exeter, 2010 [cited 16.10.12].

28. Zinman B, Gerich J, Buse JB, Lewin A, Schwartz S, Raskin P, *et al.*: **Efficacy and Safety of the Human Glucagon-Like Peptide-1 Analog Liraglutide in Combination With Metformin and Thiazolidinedione in Patients With Type 2 Diabetes (LEAD-4 Met+TZD).** *Diabetes Care* 2009,32:1224-1230.
29. Kay J, Matteson EL, Dasgupta B, Nash P, Durez P, Hall S, *et al.*: **Golimumab in patients with active rheumatoid arthritis despite treatment with methotrexate: a randomized, double-blind, placebo-controlled, dose-ranging study.** *Arthritis Rheum* 2008,58:964-75.
30. Keystone EC, Genovese MC, Klareskog L, Hsia EC, Hall ST, Miranda PC, *et al.*: **Golimumab, a human antibody to tumour necrosis factor {alpha} given by monthly subcutaneous injections, in active rheumatoid arthritis despite methotrexate therapy: the GO-FORWARD Study.** *Ann Rheum Dis* 2009,68:789-96.
31. Smolen JS, Kay J, Doyle MK, Landewé R, Matteson EL, Wollenhaupt J, *et al.*: **Golimumab in patients with active rheumatoid arthritis after treatment with tumour necrosis factor ? inhibitors (GO-AFTER study): a multicentre, randomised, double-blind, placebo-controlled, phase III trial.** *The Lancet* 2009,374:210-221.
32. Meads C, Jit M, Tsourapas A, Ashfaq K, Connock M, Fry-Smith A, *et al.*: **Tocilizumab for the treatment of rheumatoid arthritis: a single technology appraisal** *[Internet]*. Birmingham: West Midlands, Health Technology Assessment Collaboration, 2009 [cited 5.9.12].
33. *CRD and CHE Technology Assessment Group:* **Fingolimod for the treatment of relapsing remitting multiple sclerosis** *[Internet]*. York: Centre for Reviews and Dissemination/Centre for Health Economics, 2011 [cited 12.9.12].
34. Higgins JPT, Green S, eds: *Cochrane handbook for systematic reviews of interventions [Internet]*. Version 5.1.0 [updated March 2011]: The Cochrane Collaboration, 2011 [cited 23.03.11]. Available from: http://www.cochrane-handbook.org/
35. Froud R, Eldridge S, Lall R, Underwood M: **Estimating the number needed to treat from continuous outcomes in randomised controlled trials: methodological challenges and worked example using data from the UK Back Pain Exercise and Manipulation (BEAM) trial.** *BMC Med Res Methodol* 2009,9:35.
36. *On the practice of dichotomization of quantitative variables.* American Psychological Association, 2002; 19-40.
37. Posada de la Paz M, Groft SC, Carey J: **The Importance of Case Reports in Advancing Scientific Knowledge of Rare Diseases.** In: *Rare Diseases Epidemiology*.Edited by: Springer Netherlands; 2010: 77-86.
38. Griggs RC, Batshaw M, Dunkle M, Gopal-Srivastava R, Kaye E, Krischer J, *et al.*: **Clinical research for rare disease: Opportunities, challenges, and solutions.** *Mol Genet Metab* 2009,96:20-26.
39. National Institute for Health and Care Excellence: **Highly Specialised Technologies programme: interim process and methods** [Internet]. *London: National Institute for Health and Care Excellence*, 2013 [cited 24.3.14].
40. National Institute for Health and Clinical Excellence. **Fingolimod for the treatment of highly active relapsing–remitting multiple sclerosis [TA254]** [Internet]. *London: National Institute for Health and Clinical Excellence*2012 [cited 12.9.12].
41. National Institute for Health and Clinical Excellence. **Tocilizumab for the treatment of rheumatoid arthritis (rapid review of technology appraisal guidance 198)** [TA247] [Internet]. *London: National Institute for Health and Clinical Excellence* 2012 [cited 10.9.12].
42. National Institute for Health and Clinical Excellence. **Liraglutide for the treatment of Type 2 Diabetes Mellitus [TA203]** [Internet]. *London: National Institute for Health and Clinical Excellence 2010* [cited 10.9.12].
43. National Institute for Health and Clinical Excellence. **Golimumab for the treatment of rheumatoid arthritis after the failure of previous disease-modifying anti-rheumatic drugs [TA225]** [Internet]. *London: National Institute for Health and Clinical Excellence* 2011 [cited 10.9.12].
44. Jackson R, Tosh J, Davis S, Wong R, Stevenson M, Stevens J. **Golimumab for the treatment of rheumatoid arthritis after failure of previous disease-modifying antirheumatic drugs: a single technology appraisal** [Internet]. *Sheffield: ScHARR, The University of Sheffield* 2010 [cited 10.9.12].
45. National Institute for Health and Clinical Excellence. **Cilostazol, naftidrofuryl oxalate, pentoxifylline and inositol nicotinate for the treatment of intermittent claudication in people with peripheral arterial disease [TA223]** [Internet*]. London: National Institute for Health and Clinical Excellence2011* [cited 11.9.12].
46. Squires H, Simpson E, Meng Y, Harnan S, Stevens J, Wong R. **Cilostazol, naftidrofuryl oxalate, pentoxifylline and inositol nicotinate for the treatment of intermittent claudication in people with peripheral arterial disease (Assessment Report)** [Internet]. *Sheffield: ScHARR, The University of Sheffield* 2010 [cited 11.9.12].
47. National Institute for Health and Clinical Excellence. **Adalimumab, etanercept, infliximab, rituximab and abatacept for the treatment of rheumatoid arthritis after the failure of a TNF inhibitor [TA195]** [Internet]. *London: National Institute for Health and Clinical Excellence2008* [cited 12.9.12].
48. Malottki K, Barton P, Tsourapas A, Uthman A, Liu Z, Routh K, *et al.* **Adalimumab, etanercept, infliximab, rituximab and abatacept for the treatment of rheumatoid arthritis after the failure of a TNF inhibitor: a systematic review and economic evaluation (Assessment Report)** [Internet]. *London: National Institute for Health and Clinical Excellence2009* [cited 12.9.12].
49. National Institute for Health and Clinical Excellence. **Donepezil, galantamine, rivastigmine and memantine for the treatment of Alzheimer's Disease: review of NICE technology appraisal guidance 111 [TA217]** [Internet]. *London: National Institute for*

Health and Clinical Excellence2011 [cited 16.10.12].
50. Cummins E, Royle P, Shyangdan D, Waugh N**. Liraglutide for the treatment of Type 2 Diabetes: a single technology appraisal** [Internet]. Aberdeen Aberdeen HTA Group2009 [cited 5.9.12].

ADDITIONAL FILE 1: SEARCH STRATEGIES

A) The EMBASE search (OvidSP) was from 1974 to 2012/Wk33 and identified 1265 records. The search was undertaken on 21.8.12

1. enzyme replacement/ or (enzyme replacement therap$ or ERT or ERTs).ti,ab,ot. (6350)
2. iduronate 2 sulfatase/ or (Idursulfase or Elaprase or (iduron$ adj2 sulfa$) or idursulfase).mp. (819)
3. 50936-59-9.rn. (496)
4. or/1-3 (6977)
5. Hunter syndrome/ or (Hunter$ adj2 (syndrome or disease or glossitis)).ti,ab,hw,ot. (1319)
6. ((mucopolysaccharidosis or mps) adj3 (II or "2")).ti,ab,hw,ot. (837)
7. ((iduronate or sulfoiduronate) adj2 sulfatase adj2 deficienc$).ti,ab,ot,hw. (65)
8. or/5-7 (1697)
9. 4 and 8 (624)
10. enzyme replacement/ or (enzyme replacement therap$ or ERT or ERTs).ti,ab,ot. (6350)
11. (Agalsidase adj2 (alpha or alfa or beta)).mp. (841)
12. alpha galactosidase/ or (melibiase or beano).mp. (2861)
13. (galactosidase adj2 (alpha or alfa)).mp. (4039)
14. agalsidase alfa/ or agalsidase beta/ or (replagal or Fabrazyme).mp. (824)
15. (104138-64-9 or 9025-35-8).rn. (1286)
16. or/10-15 (9849)
17. Fabry disease/ or (ceramide trihexosidosis or glycosphingolipid lipidosis or glycosphingolipidosis or glycosphingolipoidosis or hereditary dystopic lipidosis or mckusick 30150).ti,ab,hw,ot. (4042)
18. (Fabry$ adj3 (Anderson or disease or syndrome)).ti,ab,hw,ot. (4358)
19. (angiokeratoma adj2 (corporis or diffus$)).ti,ab,hw,ot. (290)
20. (alpha galactosidase adj2 deficienc$).ti,ab,hw,ot. (364)
21. (Ceramide adj2 trihexosidase adj2 deficienc$).ti,ab,hw,ot. (2)
22. or/17-21 (4473)
23. 16 and 22 (2571)
24. enzyme replacement/ or (enzyme replacement therap$ or ERT or ERTs).ti,ab,ot. (6350)
25. velaglucerase alfa/ or (velaglucerase alfa or velaglucerase alpha or vpriv or glucosylceramidase glycoform alpha or Imiglucerase or Velaglucerasum or Velaglucerasa).mp. (890)
26. 884604-91-5.rn. (60)
27. Imiglucerase/ or (imiglucerase or cerezym$).mp. (873)
28. (143003-46-7 or 154248-97-2).rn. (771)
29. taliglucerase alfa/ or (taliglucerase alfa or taliglucerase alpha or uplyso or elelyso).mp. (46)
30. or/24-29 (6756)
31. Gaucher disease/ or (Gaucher$ disease or cerebroside storage disease or cerebrosidosis or kerasinosis or mckusick 23090 or mckusick 23100 or morbus gaucher).ti,ab,hw,ot. (5876)
32. (Cerebroside adj2 lipidosis adj2 syndrome$).ti,ab,ot,hw. (0)
33. (glucocerebrosidase adj2 deficienc$).ti,ab,ot,hw. (179)
34. (deficienc$ adj2 (acid or glucosylceramide) adj2 (beta-glucosidase or beta glucosidase)).ti,ab,ot,hw. (17)
35. or/31-34 (5879)
36. 30 and 35 (1494)
37. enzyme replacement/ or (enzyme replacement therap$ or ERT or ERTs).ti,ab,ot. (6350)
38. Laronidase/ or (Laronidase or Aldurazyme or alronidase or bm101 or iduronidase).mp. (1094)
39. 210589-09-6.rn. (289)
40. or/37-39 (7123)
41. Hurler syndrome/ or (chondroosteodysplasia or chondroosteodystrophy or chondroosteoplasia or dysostosis multiplex or gargoylism or gargolylism or lipochondrodystrophy or mckusick 25280).ti,ab,hw,ot. (2282)
42. ((Hurler$ or helmholtz Harrington or scheie$ or iduronidase deficienc$) adj2 (disease or syndrome)).ti,ab,hw,ot. (2311)
43. ((mucopolysaccharidosis or mps) adj3 (I or "1")).ti,ab,hw,ot. (1158)
44. or/41-43 (2990)
45. 40 and 44 (969)
46. enzyme replacement/ or (enzyme replacement therap$ or ERT or ERTs).ti,ab,ot. (6350)
47. recombinant glucan 1,4 alpha glucosidase/ or (recombinant glucan or Alglucosidase alfa or Alglucosidase alpha or Myozyme or lumizyme or pompase or recombinant acid alpha glucosidase or recombinant acid maltase or recombinant human acid

alpha glucosidase).mp. (394)
48. 420794-05-0.rn. (0)
49. or/46-48 (6497)
50. glycogen storage disease type 2/ or (Pompe Disease or cardiomuscular glycogenosis or diffuse glycogenosis or alpha glucosidase deficiency syndrome or mckusick 23230 or acid maltase deficienc$ or generali?ed glycogenos?s).ti,ab,hw,ot. (2014)
51. ((glycogen storage disease or glycogenos$) adj3 (II or "2" or pompe)).ti,ab,hw,ot. (1968)
52. ((alpha-14-glucosidase or alpha 14 glucosidase) adj2 deficienc$).ti,ab,ot,hw. (0)
53. or/50-52 (2159)
54. 49 and 53 (706)
55. enzyme replacement/ or (enzyme replacement therap$ or ERT or ERTs).ti,ab,ot. (6350)
56. Galsulfase/ or (Galsulfase or Naglazyme or aryplase or recombinant arylsulfatase B or recombinant n acetylgalactosamine or Arylsulfatase B or ARSB or rhASB).mp. (648)
57. 552858-79-4.rn. (166)
58. or/55-57 (6831)
59. Maroteaux Lamy syndrome/ or (Maroteaux Lamy or Maroteaux-Lamy or Polydystrophic Dwarfism or arylsulfatase b deficiency or mckusick 25320).ti,ab,hw,ot. (668)
60. ((mucopolysaccharidosis or mps) adj3 (VI or "6")).ti,ab,hw,ot. (595)
61. or/59-60 (901)
62. 58 and 61 (408)
63. 9 or 23 or 36 or 45 or 54 or 62 (5819)
64. Random$.tw. or clinical trial$.mp. or exp health care quality/ (2884833)
65. animal/ (1794788)
66. animal experiment/ (1632836)
67. (rat or rats or mouse or mice or murine or rodent or rodents or hamster or hamsters or pig or pigs or porcine or rabbit or rabbits or animal or animals or dogs or dog or cats or cow or bovine or sheep or ovine or monkey or monkeys).mp. (5515769)
68. or/65-67 (5515769)
69. exp human/ (13812741)
70. human experiment/ (303755)
71. or/69-70 (13814176)
72. 68 not (68 and 71) (4436054)
73. 64 not 72 (2748139)
74. 63 and 73 (1419)
75. limit 74 to embase (1265)

Randomised controlled trials filter:

Wong SS, Wilczynski NL, Haynes RB. Developing optimal search strategies for detecting clinically sound treatment studies in EMBASE. Journal of the Medical Library Association 2006;94(1):41-7.

ADDITIONAL FILES 2: QUALITY ASSESSMENTS

Quality Assessment for included randomised controlled trials of ERTs

Study Name	Randomisation	Allocation concealment	Blinding:			Incomplete outcome data:			Selective outcome reporting:	Funding/comments
	Adequate?	Adequate?	Patient blind	Physician blind	Outcome assessor blind	Attrition/ exclusion reported	Appropriate analysis methods	Incomplete outcome data addressed	Results for all specified outcomes	
Muenzer 2007[17]	Yes	NR	Yes	Yes	NR	Yes	Yes	Yes	Yes	Pharma funded.
Muenzer 2006 [18]	NR	NR	Yes	Yes	NR	Yes	ITT	Yes	Yes	Double blind & randomised, but methods not described. Funded by pharma and public
Banikazemi 2007[19]	Yes	NR	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Pharma funded
Eng 2001[21]	NR	NR	Yes	Yes	NR	NR	ITT	Unclear	Yes	Double blind & open extension, public funding.
Schiffman 2006[23]	Yes	NR	NR	NR	NR	Yes	No	No	Yes	Pharma and public funding
Schiffman 2001[22]	Yes	NR	Yes	Yes	NR	Yes	Mixed	Mixed	Yes	Double blind, pharma funded.
Bierer 2006[20]	NR	NR	NR	NR	NR	NR	NR	Unclear	Yes	Pharma funded .
Hughes 2008[26]	NR	NR	Yes	Yes	NR	Yes	ITT	Yes	Yes	Double blind & randomised, but methods not described. Funded by pharma.
Moore 2002[24]	NR	NR	Yes	Yes	NR	No	PP	No	Yes	Double blind & randomised, but methods not described. Public funding.
Hajioff 2003[25]	NR	NR	NR	NR	NR	NR	Wilcoxon	No	Yes	Described as randomised but no methods. Funding NR.

Information extracted by KSR from original papers. NA=not applicable, NR=not reported, ITT= intention to treat analysis, PP= per protocol anlaysis. Trial described as double blind assumed to be blinding of patient and physician.

Quality Assessment of comparator randomised controlled trials extracted from NICE evidence reviews.								
Appraisal and data source	Study Author	Adequate sequence generation?	Allocation concealment?	Blinding?	Incomplete outcome data addressed?	Free of selective reporting?	Groups comparable at baseline?	Sample size calculation?
TA203[50]	LEAD 4	Yes	Yes	Yes	Yes	Yes	Yes	Yes
TA225[44]	GO FORWARD	Yes	Yes	Yes	Yes	Yes	Yes	NR
	GO AFTER	Yes	Yes	Yes	Yes	Yes	Yes	NR
	Kay 2008	Yes	Yes	Yes	Yes	Yes	Yes	NR
TA223[46]	Spengel 2002	Yes	Unclear	Yes	Yes	Yes	NR	NR
TA195[48]	Malottki 2009	Yes	Unclear	Yes	Yes	NR	NR	NR
	Cohen 2006	Yes	Yes	Yes	Unclear	NR	NR	NR
TA254[33]	Kappos 2010	Yes	Yes	Yes	Yes	Yes	No	Yes
TA217 [27,49]	Rockwood 2006	Yes	Yes	Partial	Yes	NR	No	NR
	Nunez 2003	No	No	Partial	No	NR	Yes	NR
	Rogers 1998	Yes	No	Partial	No	NR	Yes	NR
	Burns 1999	No	No	Partial	No	NR	Yes	NR
	Wilkinson 2001	Yes	Yes	Partial	No	NR	Yes	NR
	Raskind 2000	Yes	NR	Partial	NR	NR	Yes	NR
	Tarlot 2000	NR	Yes	Partial	NR	NR	Yes	NR
	Wilcock 2000	Yes	Yes	Yes	Yes	NR	Yes	NR
	Brodarty 2005	Yes	Yes	Yes	Yes	NR	Yes	NR
	Corey-Bloom 1998	Yes	Yes	Partial	Yes	NR	Yes	NR
	Rosler 1999	NR	Yes	Partial	Yes	NR	Nr	NR
	Feldman & lane 2007	NR	Unclear	Partial	Yes	NR	Yes	Yes
	Reisberg 2003	NR	Yes	Partial	NR	NR	Yes	NR
	Van Dyck 2007	Unclear	Unclear	Partial	Parital	NR	Yes	NR
TA198 Meads 2009[32]	No quality tables presented.							

Information extracted by KSR from assessments prepared in evidence reviews. NA=not applicable, NR=not

ADDITIONAL FILES 3: NNTB ANALYSIS OF THE MAIN OUTCOMES

NNTB analysis of the main outcomes for ERT studies				
Source	**Outcome & follow-up**	**Treatment**	**NNTB**	**95% CI**
Hunter Syndrome:				
Muenzer et al 2007[17]	ǂUrinary GAG excretion (µg GAG/mg urine creatinine) at 6 m	Idursulfase 0.15mg/kg every other week	2.6	NNTH 5.5 to ∞ to NNTB 1.0
		Idursulfase 0.5mg/kg every other week	2.8	NNTH 4.1 to ∞ to NNTB 1.0
		Idursulfase 1.5mg/kg every other week	2.5	NNTH 6.5 to ∞ to NNTB 1.0
Muenzer et al 2006[18]	ǂUrinary GAG excretion 12 m change from baseline	Idursulfase 0.5 mg/kg weekly	1.4	1.1 to 1.7
		Idursulfase 0.5 mg/kg every other week	1.4	1.1 to 1.7
Fabry Disease:				
Hughes 2008[26]	ǂPlasma Gb3 6 m change from baseline	agalsidase alpha	17.2	NNTH 9.6 to ∞ to NNTB 4.5
Schiffman 2001[22]	ǂPlasma Gb3 6 m change from baseline	agalsidase alpha	3.2	1.7 to 23.4
Banikazemi 2007[19]	*Composite (death, renal, cardiac or cerebrovascular event), 35m	agalsidase beta	6.9	NNTH 14.8 to ∞ to NNTB 2.8
Eng 2001[21]	*Free of microvascular endothelial deposits of globotriaosylceramide, 5m	agalsidase beta	1.5	1.2 to 1.9
Bierer 2006[20]	ǂChange in VO2 max, 18m	agalsidase alpha	2.7	NNTH 2.3 to ∞ to NNTB 0.9
Moore 2002[24]	ǂChange in regional cerebral flow during visual activation, 6m	agalsidase alpha	3.4	NNTH 13.4 to ∞ to NNTB 1.5

** dichotomous outcomes; ǂ continuous outcomes; m= months; NNTB = The number needed to treat for an additional beneficial outcome; CI= confidence intervals; Gb3= globotriaosylceramide; VO2= Max oxygen uptake at peak exercise; GAG= glycosaminoglycan. Note that for NNTB the value is not always contained within the confidence intervals; this is due to calculating NNTB as 1/absolute risk difference, therefore 1/0 = ∞.*

ADDITIONAL FILES 3: NNTB ANALYSIS OF THE MAIN OUTCOMES

NNTB analysis of the main outcomes for non-ERT/ comparator studies					
Source	**Trial**	**Outcome (follow-up)**	**Treatment**	**NNTB**	**95% CI**
TA195 Rheumatoid arthritis after failure of a TNF inhibitor:					
Appraisal: Malottki 2009[48]	REFLEX	*ACR 20 mod ITT (24wks)	Rituximab 2000mg + MTX	3.0	2.3 to 4.2
	ATTAIN	*ACR 20 mod. ITT (3m)	Abatacept 10mg/kg	3.6	2.7 to 5.3
		*ACR 20 mod ITT (6m)		3.1	2.4 to 4.3
TA225 Rheumatoid arthritis:					
**GO-FORWARD[30]	**GO-FORWARD[34]	*ACR 20, ITT (14wks)	Golimumab 50mg + MTX	4.6	2.9 to 11.2
			Golimumab 100mg + MTX	4.3	2.8 to 9.9
		*ACR 20, ITT (24wks)	Golimumab 50mg + MTX	3.8	2.5 to 7.4
			Golimumab 100mg + MTX	3.8	2.5 to 7.4
**Kay 2008[29]	**Kay 2008[35]	*ACR 20, ITT (16wks)	Golimumab 50mg + MTX every 4 wks	4.4	2.2 to 1648.8
			Golimumab 50mg + MTX every 2/ 4 wks	7.8	NNTH 9.7 to ∞ to NNTB 2.8
			Golimumab 100mg + MTX every 4 wks	5.3	NNTH 22.8 to ∞ to NNTB 2.4
			Golimumab 100mg + MTX every 2/ 4 wks	2.4	1.6 to 4.7
**GO-AFTER [31]	**GO-AFTER [36]	*ACR 20, ITT (14wks)	Golimumab 50mg every 4 wks	45.2	NNTH 11.4 to ∞ to NNTB 7.6
			Golimumab 100mg every 4 wks	20.7	NNTH 16.0 to ∞ to NNTB 6.3
TA254 Relapsing remitting multiple sclerosis:					
Appraisal 2011[33]	FREEDOMS	*Absence of disability progression ITT (3m)	Fingolimod 0.5mg/day	13.6	8.2 to 39.3
		*Absence of disability progression ITT (6m)		79.8	NNTH 25.2 to ∞ to NNTB 15.5
TA203 Type 2 Diabetes:					
**LEAD 4 [28]	**LEAD 4 [37]	*HbA1c <7%, LOCF (26 wks)	Liraglutide 1.2 mg + metformin + rosigliazone 4mg	3.3	2.5 to 5.0
TA223 PAD:					
Appraisal: Squires 2010[46]	Spengel 2002	ǂPain free walking distance, ITT (24wks)	Naftidrofuryl 600mg	7.3	4.9 to 14.8
TA198 Rheumatoid arthritis:					

NNTB analysis of the main outcomes for non-ERT/ comparator studies					
Source	**Trial**	**Outcome (follow-up)**	**Treatment**	**NNTB**	**95% CI**
Appraisal: Meads 2010 [32]	OPTION	*ACR20 ITT(24wks)	Tocilizumab 8mg/kg	3.1	2.4 to 4.4
	LITHE			3.4	2.8 to 4.4
	TOWARD			2.8	2.4 to 3.2
	LITHE	*ACR20 ITT (52wks)		3.5	2.8 to 4.5
TA217 Alzheimer's Disease:					
Appraisal: Bond 2010[27]	Rogers 1998	ǂADAS (mean change from baseline), LOCF, 12 wks	Donepezil 10mg/d	4.5	3.0 to 8.7
	Nunez 2003			330.8	NNTH 7.3 to ∞ to NNTB 6.9
	Rogers 1998	ǂADAS (mean change from baseline), LOCF, 24 wks		5.4	3.4 to 13.4
	Burns 1999			4.4	3.3 to 6.8
	Wilkinson 2001	ǂADAS (mean change from baseline), LOCF, 12-16 wks	Galantamine ≤24mg/d	7.2	3.7 to 297
	Brodaty 2005			5.5	4 to 8.8
	Rockwood 2006			7.1	NNTH 32.7 to ∞ to NNTB 3.2
	Rashkind 2000	ǂADAS (mean change from baseline), LOCF, 21-26 wks		3.8	2.8 to 5.8
	Tariot 2000			5.6	4.0 to 9.4
	Wilcock 2000			5.2	3.5 to 10.1
	Brodaty 2005			5.5	4.0 to 8.7
	Corey Bloom 1998	ǂADAS (mean change from baseline), LOCF, 24-26wks	Rivastigmine ≥12mg/d	4.3	3.2 to 6.8
	Rosler 1999			10.7	5.5 to 213
	Feldman 2007			8.1	5.0 to 21.8
	Reisberg 2003	ǂSIB (mean change from baseline), LOCF, 24-28wks	Memantine 5-20mg/d	-7.4	-4 to -44.4
	Van Dyck 2007			-61.8	NNTH 8.2 to ∞ to NNTB 11.1

** dichotomous outcomes; ǂ continuous outcomes **= original trial paper used for data source. NNTB = The number needed to treat for an additional beneficial outcome (calculated in comparison to placebo). MTX =methotrexate, CI= confidence intervals; wks= weeks, m= months; d= days; LOCF=last observation carried forward; ITT=intention to treat analysis; mod ITT = modified ITT; ACR (20)=American College of Rheumatology score (patients have ≥ 20% fewer tender joints and ≥ 20% fewer swollen joints); HbA1c = glycated haemoglobin; SIB= severe impairment battery score; ADAS = Alzheimer's Disease Assessment Scale score. The results from TA217 show multiple recommended drugs with different outcomes and for each there were multiple trial results shown on separate lines. Note that for NNTB the value is not always contained within the confidence intervals; this is due to calculating NNTB as 1/absolute risk difference, therefore 1/0 = ∞.*