EEPRU

Economic Evaluation of Health and Social Care Interventions Policy Research Unit

RESEARCH REPORT

Supporting the routine collection of patient reported outcome measures in the National Clinical Audits for assessing cost-effectiveness

Work Package 1
What patient reported outcome measures should be used in the 13 health conditions specified in the 2013/14 National Clinical Audit programme?

Authors: Roberta Ara, Ana Duarte, Sue Harnan, Jo Leaviss, Steve Palmer, Mark Sculpher, John Brazier

Correspondence to: Roberta Ara, HEDS, ScHARR, University of Sheffield, Regent Court, 30 Regent Street, Sheffield, S1 4DA. Email r.m.ara@sheffield.ac.uk

RR0029

March 2015

The University Of Sheffield.

UNIVERSITY of York

The Policy Research Unit in Economic Evaluation of Health and Care interventions is funded by the Department of Health Policy Research Programme. It is a collaboration between researchers from the University of Sheffield and the University of York.

The Department of Health's Policy Research Unit in Economic Evaluation of Health and Care Interventions is a 7 year programme of work that started in January 2011. The unit is led by Professor John Brazier (Director, University of Sheffield) and Professor Mark Sculpher (Deputy Director, University of York) with the aim of assisting policy makers in the Department of Health to improve the allocation of resources in health and social care.

This is an independent report commissioned and funded by the Policy Research Programme in the Department of Health. The views expressed are not necessarily those of the Department.

EXECUTIVE SUMMARY

OBJECTIVES: To assess the appropriateness of the EQ-5D in 13 specified conditions (irritable bowel disease (IBD), epilepsy, diabetes, bowel cancer, head and neck cancer, psychological therapies, schizophrenia, dementia, cardiac arrhythmia, heart failure, coronary angioplasty, acute coronary syndrome, adult cardiac surgery); identify what measures may be used when the EQ-5D is not appropriate; identify if fields in corresponding National Clinical Audits (NCA) will suffice to conduct economic evaluations; present recommendations and future research questions in this area.

METHOD: For each condition, three reviews were considered (WP1.1, 1.2 and 1.3). WP1.1 comprised a review of the appropriateness of the EQ-5D. Known reviews were assessed in the first instance for adequacy according to a set of criteria and expert opinion. Where a known review was not available or not adequate, a systematic review of systematic reviews was conducted. Where no adequate systematic review was identified, a systematic review of primary studies was conducted. For WP1.1, systematic searches were conducted in PubMed and Embase and two conference proceedings citation indexes (Web of Science and EuroQoL) using appropriate free-text keywords and MeSH/thesaurus terms, and applying appropriate limits according to the review being conducted (review of reviews or of primary studies). WP1.2 comprised a review of condition specific measures that could be used and was only conducted for conditions where there was evidence that the EQ-5D was not adequate. This review consulted several online sources, such as Royal College websites, to identify research guidelines describing which condition specific measures should be used in research. WP1.3 comprised a review of existing cost-effectiveness evaluations used in recent health technology assessments (HTA) across the 13 conditions. Multiple technology appraisals (MTA) and single technology appraisals (STA) were systematically identified from the National Institute for Health and Care Excellence (NICE) HTA programme and the methods used to assess the cost effectiveness were extracted and compared to the data collected routinely in the NCAs.

In all reviews, retrieved citations were considered for inclusion against a set of standardised selection criteria by one reviewer, and a second reviewer consulted in cases of uncertainty. Syntheses included tabulation of study results and either: presentation of a structured abstract and critique in the case of included systematic reviews or HTAs (WP1.1, 1.3); a narrative synthesis and discussion of results in the case of primary studies (WP1.1); or a tabulation and narrative synthesis in the case of research guidelines (WP1.2).

RESULTS:

Patient reported outcome measures (WP1.1 and 1.2)

For WP1.1, conclusions were drawn from known reviews in four cases (psychological therapies; diabetes; epilepsy children; dementia); from updates and reanalyses of existing reviews in five cases (epilepsy adults; bowel cancer; head and neck cancer; schizophrenia; CVD); and from systematic reviews of primary studies in two cases (IBD adults; IBD children). For WP1.2, six reviews of the literature, augmented with literature known to the authors, were used to identify alternative or additional patient reported outcome measures (PROMs) for patients with epilepsy (paediatrics), IBD (paediatrics), schizophrenia, head and neck cancer, dementia, and patients receiving psychological therapies.

The psychometric properties of the EQ-5D were found to be adequate in ten of the 13 conditions. The exceptions were epilepsy (where the Paediatric quality of life inventory (PedsQL)^{TM} was recommended), schizophrenia (where WEMWBS, to be replaced by ReQOL, was recommended) and dementia (where DEMQOL-U was recommended).

With the exception of the mental health condition, no reference was made to the possibility of the EQ-5D being less responsive in older age groups. The presence of comorbidities in subjects was not reported in either the existing reviews or most of the primary studies. However, as the presence of comorbidities was not identified as an explicit exclusion criterion in all but the metal health review and in the diabetes studies, and the broad age ranges covered, it is assumed that the results and conclusions drawn will generalise across subjects with comorbidities.

Evidence for use in economic evaluations (WP1.3)

While the evidence collected in the individual audits will allow comparison of providers in many conditions, it is clear that the mandatory fields in most of the audits will not provide sufficient detailed information to perform formal economic evaluations. However, many of the audits contain optional fields which would be useful for economic evaluations. A recurrent issue relates to the level of detail collected and the timing of the variables collected. To be useful for economic evaluations, the variables used have to be synchronised, and/or collected over periods of time to assess progression or relapse etc.

Recommendations and associated areas for future research

There are two primary recommendations, namely a) develop a condition specific patient questionnaire for inclusion in the ten NCAs where PROMs are not currently collected, and b) expand the existing patient questionnaire to include additional settings such as primary care in diabetes and patients not undergoing ablation procedures in the cardiac arrhythmias NCA. The EQ-5D-5L is recommended for inclusion in all audits except: schizophrenia (WEMWBS recommended, to be replaced with ReQOL), dementia (DEMQOL-U is recommended), and epilepsy (PedsQL^{TM} is recommended). In addition to the EQ-5D-5L, condition specific PROMs are recommended in several of the conditions (IBD, both cancers, patients receiving psychological therapies, dementia). It is also recommended that mandatory information is collected relating to condition severity (IBD, cancer, epilepsy, diabetes, and CVD), current interventions (all conditions), clinical events (all CVD conditions, diabetes) and mortality differentiated by cause (majority of conditions), and evidence relating to the side effects associated with pharmaceutical interventions and surgical/radiotherapy complications (all relevant conditions).

Many of the suggestions for future research are generic across the conditions:

- assess the psychometric properties of the recommended PROMs using the data collected in the respective audits (*all audits*)
- synchronise the timing of collection of clinical and PROMs evidence to enable these data to be used in economic evaluations (*all audits*)
- conduct a detailed review of data collected in the current audits with a view to informing which variables should be collected as mandatory evidence to inform future economic evaluations (*all audits*)

Individual condition specific recommendations and areas for future research include:

- Conduct mapping functions between clinical variables/PROMs and preference-based measures to enable the data to be used in economic evaluations (*epilepsy*, *IBD, diabetes, bowel cancer)*
- Generate new preference-based weights for EQ-5D vision bolt-on (*diabetes*)
- Explore the effects of hypoglycaemic events on health related quality of life (HRQoL) (*diabetes*)
- Ensure all therapists use a common set of measures, which match those adopted in the National Health Service (NHS) Outcomes framework (*psychological therapies*)

INDEX

APPENDIX

TABLES

Acronym	Definition
AEs	Adverse events
BAI	Beck Anxiety Inventory
BCVA	Best corrected visual activity
BDI	Becks depression index
BMI	Body mass index (kg/m^2)
CAI	Clinical activity index
CD	Crohn's disease
CDAI	Crohn's disease activity index
CHU-9D	Child Health Utility 9D
CORE-OM	Clinical Outcomes in Routine Evaluation — Outcome Measure
DH	Department of Health
DHP	Diabetes health profile
EQ-5D	EuroQol 5 dimensions
EQ-5D-Y	EuroQol 5 dimensions youth version
ERG	Evidence review group
FR	Future research
GAD-7	Generalised Anxiety Disorder Assessment-7
HADS	Hospital anxiety and depression scale
HbA_{1c}	Glycated haemoglobin
HRQoL	Health related quality of life
HTA	Health technology assessment
HUI	Health Utility Index
HUI2	Health Utility Index mark 2
IBD	Inflammatory bowel disease
IBDQ	Inflammatory bowel disease questionnaire
MTA	Multiple technology assessment
NCA	National Clinical Audit
NHS	National Health Service
NHSSS	National hospital seizure severity score
NICE	National Institute for Health and Care Excellence
PedsQL™	Paediatric quality of life inventory™
PHQ-9	Patient health questionnaire
PR	Potential recommendations
PREMs	Patient reported experience measures
PROMs	Patient reported outcome measures
QALYs	Quality adjusted life years
QOLIE-AD-48	Quality of life in epilepsy for adolescents -48
QOLIE-89	Quality of life in epilepsy -89
R&D	Research and development
RR	Relative risk
SBP	Systolic blood pressure
SF-6D	Short form 6 dimensions
STA	Single technology assessment
TA	Technology Appraisal
UC	Ulcerative colitis

UK	United Kingdom
UKPDS	United Kingdom Prospective Diabetes Study
WP	Work package

1. BACKGROUND

EEPRU was approached by Jason Cox (Research & Development (R&D) Division) to prepare a programme of research to support the appropriateness of, and use of, patient reported outcome measures (PROMs) collected for the National Clinical Audit (NCA). The EEPRU programme was informed by a R&D template prepared by Simon Bennett, Steve Fairman and Keith Willett at National Health Service (NHS) England.

The purpose of introducing PROMs into the NCA programme is to be able to 1) compare performance between providers and commissioners in the NHS, 2) compare the cost-effectiveness of alternative providers in delivering the specific services (i.e. linking outcomes and resource use), and 3) assess the cost-effectiveness of alternative interventions and other changes in the NHS. The intention is to introduce PROMs across a range of conditions over the next 3 years commencing with 13 conditions in the 2014/15 NCA programme.

The agreed research programme consists of 3 concurrent work packages (WP) as described in the document submitted to the Depeartment of Health (DH) (8th November 2013). The current document provides an overview of the research conducted for Work Package 1 (WP1): to determine what PROMS should be used in the 13 health conditions specified in the 2014/15 NCA programme. Additional details of the methodologies used and the results for the individual conditions are provided in Appendices.

2. OVERVIEW

WP1 is split into three separate components consisting of:

- WP1.1 To examine whether the EQ-5D is appropriate in the 13 health conditions specified in the 2013/14 NCA programme.
- WP1.2 To identify what measure could be used when the EQ-5D is not appropriate in the 13 health conditions, taking into account that the proposed measure would be used to generate preference-based utility measures (either directly through existing preference-based weights, or indirectly through existing mapping functions suitable for the proposed measure).
- WP1.3 To identify the evidence required to address questions of cost-effectiveness using the NCA data.

Each component consists of a series of reviews of the literature. A synopsis of the research is provided in the following sections with more detail on the individual review objectives, methodologies, and results provided in Appendices.

3. METHOD

For each condition, three reviews were considered (WP1.1, 1.2 and 1.3). WP1.1 comprised a review of the appropriateness of the EQ-5D. In the first instance, a group of reviews known to the research team were assessed for adequacy according to a set of criteria and expert opinion (see Appendix A for more detail). For any given condition, where a known review was not available or not adequate, a recent systematic review of systematic reviews (Longworth et al. 2014b(1)) was consulted to identify other candidate systematic reviews for assessment of adequacy. Where no adequate systematic review was identified, a systematic review of primary studies was conducted.

WP1.2 comprised a review of condition specific measures and was only conducted for conditions where there was evidence that the EQ-5D was not adequate. This review consulted several online sources, such as Royal College websites, to identify research guidelines describing which condition specific measures should be used. The EQ-5D is designed for use in adults and it is not necessarily appropriate for paediatric populations. As the epilepsy NCA collects data on paediatrics only, and the inflammatory bowel disease (IBD) NCA is likely to have a relatively large proportion of paediatric subjects, additional searches were conducted to identify evidence on the psychometric properties of known paediatric generic measures: EuroQol-5D youth version (EQ-5D-Y); child health utility 9D (CHU-9D); health utility index 2 (HUI2); and paediatric quality of life inventory (PedsQL)^{TM} for these two conditions.(2-5)

WP1.3 comprised a review of existing cost-effectiveness evaluations used in recent health technology assessments (HTA) across the 13 conditions. Multiple technology appraisals (MTA) and single technology appraisals (STA) were systematically identified from the National Institute for Health and Care Excellence (NICE) HTA programme and the methods used to assess the cost-effectiveness were extracted and compared to the data collected routinely in the National Clinical Audits.

A summary of the results is provided below with additional detailed results for the individual conditions provided in Appendices C-K.

4. RESULTS FOR INFLAMMATORY BOWEL DISEASE

4.1 Evidence of appropriateness of EQ-5D in Inflammatory Bowel Disease (IBD) (WP1.1)

As no existing review was known or identified, a review of primary studies was conducted. The searches identified just two studies (N=152;(6) N=502(7)) which satisfied the inlusion criteria.(6)-23) Both studies involved adults (≥17 years old) with either Crohn's disease (CD) or ulcerative colitis (UC). Both were conducted in Germany and just one used the required United Kingdom (UK) preference-based EQ-5D index.

The EQ-5D was well accepted with less than 3.5% of EQ-5D responses missing (compared to 9% for the irritable bowel disease questionnaire (IBDQ)). Althought there was some evidence of a potential ceiling effect, construct validity was generally good on all EQ-5D health dimensions (except self-care in patients with UC) when assessed against the disease activity indices, and when discriminating between those with active disease and those in remission. Some problems were observed in the health dimensions pain/discomfort and anxiety/depression for outpatients compared to inpatients. There was some evidence of the EQ-5D being a responsive measure in those with active disease (further away from full health) who reported a deterioration in health, but there may be problems with the EQ-5D's responsiveness when detecting changes in less severely ill patients and those in remission. However, small numbers of patients in the sub-group analyses made it difficult to draw robust conclusions. In conclusion, the evidence supporting the appropriateness of the EQ-5D in adults with IBD is limited but fair but there is a lack of evidence relating to how appropriate the EQ-5D is for surgical health states. Evidence is required in paediatrics, and in patients with more severe disease, both of which are relevant for the IBD NCA population (Table 1).

4.2 Evidence of alternative measures in IBD (WP1.2)

Searches identified no studies presenting evidence relating to the psychometric properties of the four pre-specified preference-based measures in paediatrics with IBD. However, eight articles, representing seven studies, which examined the psychometric properties of alternative measures were identified and reviewed.(8-15) While primary research in this area appears to be growing, with evidence of development of several PROMs targeted at paediatrics with IBD, the evidence identified which could be used to compare PROMs in this population was limited. The searches did not identify any evidence which could be used to generate quality adjusted life years (QALYs) in this population. The most likely target measures for inclusion in the IBD NCA are the IMPACT-III and the PedsQL^{TM} v4. Based on the evidence reviewed, the target age group, and alternative responder versions, the PedsQL^{TM} is recommended over IMPACT-III measure (Table 1). However, research is

required to generate an associated preference-based tariff for the PedsQL^{TM} (or a mapping function to one of the alternative preference-based generic measures) which could then be used to generate utility values for use in cost-effectiveness evaluations.

Table 1: Summary of evidence on PROMs for IBD

Measure (N)	Target Age (years)	Target Responder	Acceptability	Reliability	Construct (KGV; Convergent)	Responsiveness (Change over time; Ceiling effects)
Adults						
EQ-5D (2)	-	-	Good	Good	Good; Good	Mixed; Mixed
Acceptable but requires additional validation (n studies =2) particularly in patients with severe IBD and those undergoing surgical procedures. Is not appropriate for paediatrics with IBD.						
PRO2, PRO3 (1)	-	-	Very limited evidence available (n studies =1)			
Acceptable but requires additional validation and is only suitable for adults with CD (not for UC).						
Paediatrics						
IMPACT-III (3)	≥9	SR	Good	Good	Mixed; Good	No evidence
Acceptable but requires additional validation (n studies = 3) and cannot be used to generate QALYs						
PedsQL (5)	2-18	SR;PR	Good	Good	Good	No evidence
Acceptable (n studies = 5) but cannot be used to generate QALYs						
PedsQL GI module (1)			Good	Good	Good; No evidence	No evidence; poor
Acceptable, but very limited evidence (n studies = 1), would need to be used as an adjunct to the PedsQL core measure, and cannot be used to generate QALYs						

N=number of studies; KGV=known group validity; CE=ceiling effect; SR=self-report, PR=Parent/carer-report

4.3 Evidence for economic evaluations in IBD (WP1.3)

The recent HTAs for IBD evaluated interventions which are relevant to those provided to patients in the IBD NCA which gathers information from secondary care settings. One compared several pharmaceutical interventions for the treatment of moderate to severe CD or fistulising CD in both adult and paediatric populations,(16) and the second compared surgery (colectomy) with rescue therapy (standard care or alternative pharmaceutical interventions) for avoidance or delay of surgery and symptom free remission in hospitalised patients with acute exacerbation of UC.(17) The key clinical information required to inform a standard economic evaluation comparing interventions (either anti-TNF agents, or surgical procedures) in IBD in the secondary care setting would be: condition severity, therapy regimens, the associated remission, relapse and withdrawal rates, the rates and types of surgical interventions and complications, and preference-based utility values. While the current IBD audit collects some information on the majority of these areas, there are some obvious omissions (see Appendix C). One of the key issues is the timing of the data collection. As IBD is characterised by periods of 'flares' and remission; for the data to be useful for economic modelling purposes, the collection needs to be synchronised in terms of timing.

4.4 Recommendations for IBD

In summary, the EQ-5D appears to be appropriate in adults with IBD, the PedsQL appears to be the most appropriate measure for paediatrics (although there are limitations with the usefulness of this measure for economic evaluations), and the current IBD audit collects much of the information required to conduct economic evaluations. However there are caveats associated with these conclusions which require consideration. The issues and corresponding potential recommendations (PR) and areas for future research (FR) are tabulated below (Table 2) with the corresponding narrative provided in Appendix C.

Table 2: Recommendations and associated future research for IBD

PR.1	*Include the new version of the EQ-5D (EQ-5D-5L) in future adult patient questionnaires*
FR.1	*Assess the psychometric properties of the EQ-5D-5L in adults with IBD using data collected in the audit*
PR.2	*Include the PedsQL™ (and the PedsQL™ GI module) in future paediatric patient questionnaires*
FR.2	*Investigate potential collaboration with the developers of the PedsQL™ with a view to developing a methodology to generate preference-based utility measures directly or indirectly (via mapping to alternative measure) from the PedsQL™.*
PR.3	*Include age related paediatric preference-based health related quality of life (HRQoL) instrument (e.g. CHU-9D, HUI2 and EQ-5D-Y) in future paediatric patient questionnaires*
FR.3	*Assess the psychometric properties of the paediatric preference-based tools in IBD using data collected in the audit*
PR.4	*Synchronise the timing of collection of a clinical measure (such as the Crohn's disease activity index (CDAI) for patients with CD, or the clinical activity index (CAI) for patients with UC) and the HRQoL measure*
FR.4	*Conduct analyses to generate mapping functions between the suggested clinical and preference-based measures to enable the evidence to be used in economic models*
PR.5	*Include an additional PROM to capture disease severity, such as the PRO2 or PRO3 (and equivalent measures for UC and paediatrics), in the patient questionnaire*
FR.5	*Assess the validity of the PRO2/PRO3 using data collected in the audit.*
FR.6	*Conduct research to generate equivalent condition severity PROMs in adults and paediatrics with UC.*
PR.6	*Include additional mandatory fields in the IBD audit such as response to current treatment, relapse and current disease activity (linked by time to HRQoL)*
FR.7	*Detailed analyses of fields currently collected in the IBD audit to identify recommendations for future mandatory fields*
PR.7	*Utilise links between the IBD audit and the new IBD register*

5. RESULTS FOR EPILEPSY

5.1 Evidence of appropriateness of EQ-5D in epilepsy (WP1.1)

The epilepsy NCA is in paediatrics, consequently the EQ-5D, which is recommended for adults, is potentially not relevant for this population. However, evidence from studies in adults will provide a useful indicator of the appropriateness of the EQ-5D in a younger population. An existing systematic review was updated and a total of five studies (reported in seven publications) were included in the current review.(18-24) The evidence base assessing the psychometric properties of the EQ-5D in adults with epilepsy on the whole is positive albeit limited. There is some evidence of a ceiling effect compared to other generic measures (SF-6D and health utility index (HUI)).(23) Support for construct validity was relatively good with evidence the EQ-5D was able to detect differences in sub-groups categorised by seizure frequency,(19) and the National hospital seizure severity score (NHSSS),(88) despite the correlations between the EQ-5D and both the severity and control of seizures being small.(23) While there was also no correlation between responses on the EQ-5D health dimensions and presence of comorbidity in one study,(23) a relationship between the EQ-5D anxiety/depression health dimension and hospital anxiety and depression scale (HADS) scores was reported.(24) There was also evidence to suggest the EQ-5D was able to detect changes in HRQoL over time in patients sub-grouped by achievement of seizure reductions,(21) or improvement or deterioration on the global rating of change.(23) However, one study suggested the magnitude of change in mean EQ-5D was smaller than might have been expected based on changes in the quality of life in epilepsy-89 (QOLIE-89) measure,(23) and in another study improvements in EQ-5D scores were not statistically significant in patients pre- versus post-surgery, where other measures were.(20) In conclusion, the evidence suggests that the EQ-5D may not be the most appropriate measure in adults, though it may be regarded as adequate. However, the EQ-5D is not relevant for paediatrics, the subject of the epilepsy NCA.

5.2 Alternative measures in epilepsy (WP1.2)

Although limited, the evidence suggests the EQ-5D is adequate for adults, inferring the youth version is worth considering for paediatrics. However, the EQ-5D-Y is only suitable for older children and no evidence was identified on its appropriateness in paediatrics with epilepsy. The evidence identified which could be used to compare PROMs in this population was limited. The searches, although limited in scope due to the time constraints of the project, did not identify any evidence which could be used to generate QALYs directly from PROMs in this population. The most likely candidate measures for inclusion in the epilepsy NCA are the Quality of life in epilepsy for adolescents -48 (QOLIE-AD-48) and the PedsQL^{TM} v4. Based on the evidence reviewed, the target age group, and

alternative responder versions, the PedsQL™ is recommended over QOLIE-AD-48 measure (Table 3). However, research is required to generate an associated preference-based tariff for the PedsQL™ (or a mapping function to one of the alternative preference-based generic measures) which could be used to generate utility values for use in cost-effectiveness evaluations.

Table 3: Summary of evidence on PROMs for epilepsy

Measure (N)	Target Age (years)	Target Responder	Acceptability	Reliability	Construct (KGV; Convergent)	Responsiveness (Change over time; Ceiling effects)
Adults						
EQ-5D (5)	-	-	No evidence	No evidence	Good; Good	Mixed; Poor
Adequate but the evidence on the different psychometric properties is limited (n studies =5) Is not appropriate for paediatrics with epilepsy.						
Paediatrics						
QOLIE-AD-48 (3)	11-17	SR	No evidence	Good	Good; Mixed	No evidence; Good
Adequate but evidence is limited (n= 3 studies and not all properties tested) Would require a systematic literature review to identify additional evidence. Cannot be used to generate QALYs and only appropriate for adolescents.						
PedsQL (5)	2-18	SR;PR	No evidence	Good	Good; No evidence	Unclear; no evidence
Acceptable (n studies = 5) but cannot be used to generate QALYs						
PedsQL epilepsy module (1) Currently under development but could be considered in the future						

KGV = known group validity; N = number of studies

5.3 Evidence for economic evaluations in epilepsy (WP1.3)

The economic models used in one STA and one clinical guideline (CG) were reviewed. The STA examined the clinical and cost-effectiveness of a pharmaceutical intervention in adults with partial refractory epilepsy,(25) while the CG compared interventions in primary and secondary care for both adults and children.(26) It is worth noting that a) the results of the searches for preference-based utilities indicated that the volume of evidence in this patient group was sparse, particularly in paediatrics, and b) it was reported that utilities were a major driver of the cost-effectiveness results.(25)

In summary, the existing HTA models were constructed around the number of seizures, the reduction in numbers of seizures (response) due to treatment, withdrawal from treatment due to adverse events (AEs), and HRQoL scores categorised by current health status. The fields in the current NCA provide insufficient detailed mandatory information to examine changes in frequency of seizures, the epilepsy medications taken (and duration), or withdrawal due to AEs or non-response. The epilepsy patient questionnaire does not currently collect HRQoL information,

concentrating on patient reported experience measures (PREMs), and the mandatory fields do not include a surrogate measure which could be used to estimate proxy utility values. It is not known if surgical interventions and associated information might be considered for collection in future audits but this is evidence that would be useful for comparing providers or the cost-effectiveness of alternative policies. While it is currently not possible to compare the cost-effectiveness of interventions in epilepsy using the audit data, it may be possible to compare performance across units.

5.4 Recommendations for epilepsy

In summary, the paediatric literature review indicates the PedsQL is an appropriate measure for paediatrics with epilepsy, but it cannot currently be used to generate utility values. The results of the review of existing cost-effectiveness HTA models indicated that the evidence base of existing preference-based data in patients with epilepsy was extremely sparse, particularly in paediatrics, and that economic models are sensitive to the utility values used (see Appendix D). Finally, the current epilepsy NCA does not provide sufficient mandatory information to compare providers or the cost-effectiveness of interventions as used in general clinical practice. The issues and corresponding potential recommendations (PR) and areas for future research (FR) are tabulated below (Table 4) with the corresponding narrative provided in Appendix D.

Table 4: Recommendations and associated future research for epilepsy

PR.8	*Include the PedsQL™ (and the PedsQL™ epilepsy module) in future epilepsy audits*
PR.9	*Include age related paediatric preference-based HRQoL instrument (e.g. CHU-9D, HUI2 and EQ-5D-Y) in future paediatric audits*
FR.8	*Assess the psychometric properties of the paediatric preference-based tools in epilepsy using data collected in the audit*
PR.10	*Include additional mandatory fields in the epilepsy audit severity and frequency of seizures, response to current treatment, and medications (and if applicable, evidence relating to surgical interventions), linked by time to HRQoL.*
FR.9	*Detailed analyses of fields currently collected in the epilepsy audit to identify recommendations for future mandatory fields*

6 RESULTS FOR DIABETES

6.1 Evidence of appropriatness of EQ-5D in diabetes (WP1.1)

An existing systematic review was updated and 16 studies were included in the current review.(27;28) The majority of studies appraised the UK EQ-5D tariff, and all used adult samples. Acceptability and reliability were both reported to be good. There was some evidence of a ceiling effect in patients with diabetes, which may be more relevant in newly diagnosed patients who do not have diabetes related complications, and are thus more likely to score relatively high on the index. The majority of studies reported the construct validity of the EQ-5D was good when compared to diabetes specific clinical and quality of life measures. Exceptions included, for example, levels of visual acuity, and potentially glycated haemoglobin levels (HbA_{1c}). Poor correlations against some variables are of less concern where the comparator may not reasonably be expected to produce a correlation with HRQoL. For example, the relationship between HbA_{1C} and HRQoL is complex; HbA_{1c} levels are an indication of blood glucose levels over the previous 2-3 months, whereas the EQ-5D asks patients what their HRQoL is today. In conclusion, the EQ-5D is adequate in patients with diabetes but additional research is required before it can be recommended for patients with visual problems (Table 5).

Table 5: Summary of evidence on EQ-5D for diabetes

Measure (N)	Acceptability	Reliability	Construct (KGV; Convergent)	Responsiveness (Change over time; Ceiling effects)
Adults				
EQ-5D (16)	Good	Good	Good; Mixed	Good; Poor
Adequate, with exception of potential problems in patients with vision problems.				

6.2 Alternative measures in diabetes (WP1.2)

Based on the psychometric properties of the EQ-5D reported for patients with diabetes, with the exception of potential problems in patients with vision problems, and the suggestion that there may be a ceiling effect, the evidence suggests the EQ-5D is appropriate in adults with diabetes. Consequently the evidence on other condition-specific or generic measures was not reviewed. It is worth noting, however, that the NHS outcomes framework uses the Diabetes health profile (DHP) self-reported outcome measure in conjunction with the EQ-5D. The DHP is available in two forms, DHP-1 and DHP-18. DHP-18 takes less time to complete, is available in electronic formats as well as paper, and there is some limited evidence and ongoing research relating to its use in cost utility

analysis.(29) It aims to capture the impact of diabetes on everyday social and emotional functioning, which may not be captured by other measures.

The problems with the EQ-5D in vision have been noted elsewhere.(30;31) and a bolt-on for vision has been developed. Patients are asked to select between "I have no problem seeing"; "I have some problems seeing" and "I have extreme problems seeing". The impact of responses on the valuation of health states has been tested in an exploratory and a full valuation study.(31) The vision bolt-on has been shown to significantly impact on at least some health states, with complex interplay between severity of the vision response, and severity of responses in the other dimensions but the authors caution that further research with larger sample sizes is required.

6.3 Evidence for economic evaluations in diabetes (WP1.3)

The economic models in six STAs and one CG were reviewed. Five of the studies compared insulin,(32-36) and the remaining two STAs compared interventions for diabetic related macular oedema.(37;38) It was noted in several evidence review group (ERG) reports that the key areas of uncertainty in the models comparing insulin control were the HRQoL parameters used for hypoglycaemic events and treatment related changes in weight. In the models exploring interventions for diabetes related visual acuity, the cost-effectiveness estimates were also noted to be sensitive to the utility values used.

Diabetes is a complex disease and patients are at risk of multiple diabetes related complications. The existing HTA models comparing insulin therapies used the United Kingdom Prospective Diabetes Study (UKPDS) risk functions to model the benefits of treatments in terms of reductions in both micro and macrovascular complications. The key variables (HbA_{1c}, body mass index (BMI), systolic blood pressure (SBP)) required for the functions are noted as optional fields in the current core diabetes NCA (Appendix E). While many of the variables required are not currently listed in the NCA, they may be available from GP records via the primary care audits. The existing HTA models comparing interventions for diabetes related visual acuity use a clinical grading measure (e.g. best corrected visual activity (BCVA)) to describe health states within the model, and changes in these to represent the effectiveness of the intervention. While the optional fields contain a field relating to eye screen attendance, there is no information which suggests that the results of eye screens are recorded in the NCA, or that presence or severity of macular oedema is recorded.

No patient reported outcome measures are currently collected in the core diabetes NCA and while it is understood that there will be some data on patients' experience of services collected in subsequent audits, there are currently no fields relating to HRQoL or any alternative measure which could be used to generate the preference-based data required to inform cost-effectiveness models. However, depending on the level of responses collected in the inpatient audit, with additional fields added, it is possible that the diabetes NCA could be used to compare providers and the cost-effectiveness of interventions.

6.4 Recommendations for diabetes

In general, the EQ-5D appears to be adequate in patients with diabetes, and based on the assumption that the audit can be linked to patients' primary care records, the current IBD audit collects much of the information required to conduct economic evaluations. The exceptions in both cases are in patients with visual problems, and information relating to HbA_{1c} and hypoglycaemic events. Potential recommendations (PR) and areas for future research (FR) are tabulated below (Table 6) with the corresponding narrative provided in Appendix E.

Table 6: Recommendations and associated future research for diabetes

PR.11	*Include the new version of the EQ-5D (EQ-5D-5L) and the DHP in future adult patient questionnaires*
FR.10	*Assess the psychometric properties of the EQ-5D-5L and the DHP in adults with diabetes using data collected in the audit*
PR.12	*Include the vision bolt-on to the EQ-5D for patients with vision problems*
FR.11	*Conduct a study to generate preference-weights for the EQ-5D vision bolt-on*
PR.13	*Include a clinical measure such as the BCVA or vision acuity score in the audit (collected at the same time as the HRQoL variable)*
FR.12	*Conduct a study exploring the effect on HRQoL associated with hypoglycaemic events and the associated fear of future events using data collected in the audit*
PR.14	*Include paediatric preference-based HRQoL instruments (e.g. CHU-9D and the HUI2 or PedsQL) in future paediatric questionnaires*
FR.13	*Assess the psychometric properties of the paediatric preference-based tools in paediatrics with diabetes using data collected in the audit*
FR.14	*Detailed analyses of fields currently collected in the IBD audit to identify recommendations for future mandatory fields*

7 RESULTS FOR BOWEL CANCER

7.1 Evidence of appropriateness of EQ-5D in bowel cancer (WP1.1)

An existing systematic review was updated and nine studies were included in the current review.(31;39) All involved patients with colorectal cancer (no information on age distribution). In seven studies patients were undergoing or had recently undergone surgery,(168- 173) whilst in two studies patients received a pharmaceutical intervention.(40;41) With the exception of one study which used UK preference-based weights,(40) it is unclear which tariff was used.

No evidence was reported in the review for either acceptability or reliability of the EQ-5D. Construct validity was explored using known group methods in six studies, and convergent methods in one study. The evidence relating to construct validity by known group methods was mixed, with two studies reporting that the EQ-5D was able to detect differences between groups where other measures did,(42;43) two studies reporting that the EQ-5D failed to detect a difference in groups where other measures did,(44;45) one study reporting no differences between groups by the EQ-5D and EORTC QLQ-C30(46) and one study reporting no difference in EQ-5D scores between two groups where it was not clear if a difference should be expected.(47) The evidence-base relating to convergent validity was small with only one relatively small (n=104) study examining this. Results relating to responsiveness were also mixed, with one study showing the EQ-5D was able to detect a change where another measure, (7) two studies showing the EQ-5D was not able to detect a change where other measures did,(43;48) and two studies showing mixed results where the EQ-5D was able to detect some changes but not others.(41;46) In summary, whilst there is some strong positive evidence to support the use of the EQ-5D in this patient group, the negative and mixed evidence suggest that additional validation is required before the EQ-5D can be recommended (Table 7).

Table 7: Summary of evidence on EQ-5D for bowel cancer

Measure (N)	Acceptability	Reliability	Construct (KGV; Convergent)	Responsiveness (Change over time; Ceiling effects)
Adults				
EQ-5D (9)	Not reported	Not reported	Mixed; Poor (n=1)	Mixed; Not reported
EQ-5D requires additional validation.				

7.2 Alternative measures in bowel cancer (WP1.2)

It is recommended that a cancer specific PROM, the European Organization Quality of Life Questionnaire (EORTC QLQ-C30) and its relevant bowel-specific module EORTC CR29 is collected alongside the EQ-5D as condition specific measures may be more sensitive to the effects of interventions on the condition specific symptoms and the side effects of treatments. The EORTC

QLQ-C30 consists of 30 questions covering function (e.g. cognitive, emotional, physical, role, social) and the common cancer symptoms (e.g. fatigue, nausea and vomiting, pain).(49) Responses to these are summarised using 14 sub-scales plus a global quality of life scale.(49) A recently developed UK based preference-based utility tariff can be used to generate utility values for use in economic evaluations.(50) However, it should be noted that the the utility values obtained using this tariff are not directly comparable to those generated using the EQ-5D.

7.3 Evidence for economic evaluations in bowel cancer (WP1.3)

Economic models in six HTAs and one CG were reviewed.(51-55) The majority examined the clinical and cost-effectiveness of chemotherapy with just one study comparing laparoscopic surgery with open-resection.(56) The results of the searches conducted to inform the model parameters suggest the volume of EQ-5D data in patients with colorectal cancer is very limited with many of the authors recommending this as a future research priority. It has also been suggested that evidence categorised by Duke's stage would be useful. Staging is currently used for chemotherapy licensing indications consequently this information would be useful when modelling the cost-effectiveness of screening interventions; it could be linked to audit data to estimate resource usage by stage; and if stage was linked to EQ-5D utility values this could be used to populate future economic models in bowel cancer.

The NCA information on clinical interventions (tumour, treatment, follow-up) would provide some of the information required to compare alternative treatments. The mortality date could be used to model overall mortality, and there may be sufficient detail to extract survival curves for progression and recurrence from the mandatory fields. Side-effects and adverse events due to chemotherapy, radiotherapy and surgery are prevalent. While there is some information on toxicity (treatment related morbidity: mild toxicity, moderate toxicity, severe toxicity, death due to toxicity) this field is non-mandatory and it is not clear if the level of detail collected would suffice to populate a model.

Assuming the mandatory fields have relatively high completion rates, with the exception of HRQoL, disease severity and toxicity, the information currently collected in the existing NCA would provide the majority of information required to model the cost-effectiveness of interventions and policies in bowel cancer. As previously noted, there is a dearth of preference-based HRQoL evidence in bowel cancer, and the collection of utilities (EQ-5D) within the NCA would be recommended as an important and valuable consideration.

7.4 Recommendations for bowel cancer

The evidence base relating to the appropriateness of the EQ-5D in patients with bowel cancer is mixed and further validation is required. All the evidence is in adults and while the NCA inclusion definition includes all patients with bowel cancer, the proportion of paediatrics is likely to be extremely small and is not considered in this report. Published evidence which could be used to populate the utility values in economic models is extremely poor. With the exception of information on treatment related adverse events/complications, it is thought that the current NCA collects much of the information required to conduct economic evaluations. However, as far as we are aware, there does not appear to be a patient questionnaire. Potential recommendations (PR) and areas for future research (FR) are tabulated below (Table 8) with the corresponding narrative provided in Appendix F.

Table 8: Recommendations and associated future research for bowel cancer

PR.15	*Include the EQ-5D in future patient questionnaires alongside a condition specific measure such as the EORTC QLQ-C30 and the colorectal module (QLQ-CR29)*
FR.15	*Assess the psychometric properties of the EQ-5D using the data collected in the bowel cancer audit*
PR.16	*Include a severity measure such as Dukes' staging (to be collected alongside the HRQoL data)*
PR.17	*Collect mandatory information on adverse events associated with chemotherapy regimens (plus radiotherapy adverse events, and surgical complications)*
PR.18	*Include additional mandatory fields in the bowel cancer audit*
FR.16	*Detailed analyses of fields currently collected in the bowel cancer audit to identify recommendations for future mandatory fields*

8. RESULTS FOR HEAD AND NECK CANCER

8.1 Evidence of appropriateness of EQ-5D in head and neck cancer (WP1.1)

One review was identified which covered all cancers.(31) No studies for head and neck cancer were found by this review. The searches were conducted in August 2010(31) and an updated search was conducted for this project in May 2014. The update searches retrieved 32 citations. None of these studies met the inclusion criteria of WP1.1. While two studies were identified in patients with brain cancer, this condition is excluded from the head and neck cancer NCA hence the studies are not reviewed here. As such, there does not appear to be any evidence relating to the appropriateness of the EQ-5D in patients with head and neck cancer.

8.2 Alternative measures in head and neck cancer (WP1.2)

The evidence identified on alternative measures was extremely limited (Appendix G) and in keeping with the recommendations for the bowel cancer audit, it is recommended that the EQ-5D and the EORTC QLQ-C30 (plus the QLQ-H&N35 module) are collected in the NCA with a view to assessing the psychometric properties of the measures using the NCA data (Table 9).

Table 9: Summary of evidence on PROMs in head and neck cancer

Measure (N)	Acceptability	Reliability	Construct		Responsiveness		Overall
			KGV	Convergent	Change over time	Ceiling Effect	
EQ-5D (0)	NE	NE	NE	NE	NE	NE	No evidence
EORTC QLQ-C30 EORTC QLQ-H&N35	The psychometric properties of these measures have not been reviewed in the current report						

N= number of studies used to inform conclusions, KGV: known group validity; NE, no evidence was identified;

8.3 Evidence for economic evaluations in head and neck cancer (WP1.3)

Two STAs relating to head and neck cancer were identified from the searches.(57;58) Both examined the clinical and cost-effectiveness of a pharmaceutical intervention plus radiotherapy compared to radiotherapy alone in patients with recurrent and/or metastatic head and neck cancer,(58) or patients with locally advanced squamous cell carcinoma of the head and neck.(57) Both studies quality adjusted survival by assigning mean utility values to the discrete health states. Presumably due to the lack of more appropriate evidence, neither study used preference-based utility values obtained from patients with head and neck cancer.

The head and neck cancer audit does not include a patient questionnaire thus PROMs are not currently collected and there does not appear to be an alternative field which could be used to predict the required preference-based utility values. Assuming there is a relatively high completion rate, it is thought that this audit collects much of the information required to derive survival curves for progression and recurrence of the disease but again key information may be missing such as condition severity (measured using information of lesions and required to case-mix when comparing providers), current pharmaceutical interventions (type of intervention, concomitant medications, remission rates, relapse rates, adverse events) and surgical rates and complications.

8.4 Recommendations for head and neck cancer

The searches conducted to identify evidence on the appropriateness of the EQ-5D found no relevant studies and published evidence which could be used to populate the utility values in economic models is extremely sparse. While it is likely that with exceptions, the current head and neck cancer audit collects much of the evidence needed to perform economic evaluations, this is far from clear. In addition, the head and neck audit is completed by clinicians/NHS staff and does not currently include a patient completed component. Potential recommendations (PR) and areas for future research (FR) are provided below (Table 10) with the corresponding narrative provided in Appendix G.

Table 10: Recommendations and associated future research for head and neck cancer

PR.19	*Include a patient questionnaire or the provision for electronic collection of PROMs*
PR.20	*Include the EQ-5D in future patient questionnaires alongside a condition specific measure such as the EORTC QLQ-C30 and the QLQ-H&N35 module*
FR.17	*Assess the psychometric properties of the EQ-5D using the data collected in the head and neck cancer NCA*
PR.21	*Information on adverse events associated with chemotherapy regimens and the side effects of surgical interventions and radiotherapy*
PR.22	*Collect an appropriate measure of severity, such as mucositis grade alongside information on lesions*
PR.23	*Include additional mandatory fields in the head and neck cancer audit*
FR.18	*Detailed analyses of fields currently collected in the head and neck cancer audit to identify recommendations for future mandatory fields*

9. RESULTS FOR PSYCHOLOGICAL THERAPIES

9.1 Evidence of appropriateness of EQ-5D in psychological therapies (WP1.1)

An existing systematic review provided evidence on the appropriateness of EQ-5D in psychological therapies.(59) A total of 20 studies evaluated the construct validity or responsiveness of the EQ-5D. Studies were conducted in a wide range of countries and it is not clear in many if the required UK preference weights were used. (60-80) Mean ages ranged from 39.6 years(78) to 74.1 years.(66) Ten studies focused on individuals with depression((60;64;65;67-70;73;78;79), 3 studies focused on individuals with anxiety,(71;72;80) and 3 focused on individuals with either depression or anxiety.(61;74;75) The remaining 2 studies were surveys of the general population, aiming to identify individuals with postnatal depression (63) or depression or anxiety.(76)

The authors concluded that while the evidence base supports the use of the EQ-5D in patients with anxiety and depression, there is evidence to suggest the EQ-5D may lack responsiveness in the elderly.(59) They also noted a stronger correlation with depression scales than anxiety scales in patients with anxiety which suggests the known group validity results may be driven by the presence of comorbid depression or the depression aspect within anxiety disorders. The evidence suggests the EQ-5D is appropriate in patients with depression, but additional research is required to confirm its appropriateness in patients with anxiety conditions (Table 11).

Table 11: Summary of evidence on EQ-5D for patients receiving psychological therapies

Measure (N)	Acceptability	Reliability	Construct (KGV; Convergent)	Responsiveness (Change over time; Ceiling effects)
Adults				
EQ-5D (21)	Not reported	Not reported	Good; Good	Mixed; not reported
	Authors note that EQ-5D correlations were higher when compared against depression than anxiety scales in patients with anxiety (study n=1), that there may be a lack of responsiveness in older adults (study n=1) and that the EQ-5D showed greater changes at the lower end of the HRQoL spectrum. The EQ-5D is appropriate in patients with depression, but additional research is needed to confirm its appropriateness in patients with anxiety.			

9.2 Alternative measures in psychological therapies (WP1.2)

Evidence from WP1.1 for psychological therapies suggests that the EQ-5D is appropriate for use in depression and additional research is needed to confirm its appropriateness in patients with anxiety. The most relevant evidence for WP1.2 comes from the Oxford PROMS group.(81) This report did

not, however, give one single recommendation and did not consider all measures available. Of most relevance to WP1.2 are the report's considerations about choosing a measure that covers both anxiety and depression, for which Clinical Outcomes in Routine Evaluation - Outcome Measure (CORE-OM) is preferred over Hospital Anxiety and Depression Scale (HADS), as it has a better level of evidence to support it and includes social function. The report states that if a preference measure is required, EQ-5D is preferred over the SF-6D.

For adult populations, the Royal College of Psychiatrists give as examples the patient health questionnaire (PHQ-9), a 9-item depression scale intended to diagnose and monitor depression; the Generalised Anxiety Disorder Assessment -7 (GAD-7), a seven-item questionnaire used as a screening tool and severity measure for generalised anxiety disorder; CORE-OM, used before and after therapy. In older adults, the Royal College recommends GAD-7 and HADS, but notes that this latter may miss somatic symptoms.

It is also worth noting that a new measure is in development that is intended to be suitable for use across the spectrum of psychotic and non-psychotic mental health conditions. The measure, Recovering Quality of Life (ReQoL) is currently under development by the Policy Research Unit in Economic Evaluation of Health and Care Interventions (EEPRU) and is due to be available around July 2015.(82) Once the measure is available and has been validated in people with depression and anxiety, the ReQoL may become a candidate measure for inclusion in the NCA.

9.3 Evidence for economic evaluations in psychological therapies (WP1.3)

As no relevant TAs were identified, the most recent clinical guideline (CG) in anxiety and depression was reviewed.(83) The CG encompassed a broad decision space covering pharmacological and physical interventions, services (organisation of care, development of staff roles, introduction of mental health specialists into primary care), and psychological and psychosocial interventions. The guideline team identified four studies describing economic evaluations for either low-intensity (84;85) or high-intensity psychological interventions.(86;87) The clinical pathway was described using discrete health states based on severity of depression using well-established cut-offs relating to the Becks Depression Index (BDI). Mean EQ-5D scores were assigned to the discrete health states and the analysts reported the relationship between EQ-5D and severity of depression (measured using the BDI mapped onto the CORE-OM) was non-linear.

The NCA collects information from patients with anxiety or depression receiving psychological therapies in the community, and includes a therapist questionnaire and a service user's questionnaire. The existing models used health states categorised by severity of condition and intervention specific relapse rates were used to compare individual therapies. Responses to the type of psychological therapy provided, sub-categorised by high (e.g. cognitive analytic therapy) or low (e.g. psycho education) intensity therapy, the number of sessions attended and completion of therapy, are mandatory in the retrospective service user questionnaire. These could potentially be used to model adherence and withdrawal rates but it is not clear if there are any fields which could be used to model relapse, which is a frequent occurrence in this chronic condition and a key parameter for any economic model in this area.

The retrospective service user questionnaire also includes information on initial and final outcome scores such as HADs, CORE-10, and Beck anxiety index (BAI) which would be useful to measure severity. While in theory these could be used to identifying response to treatment, this would depend on the timings of the data collection. There is evidence in the literature which could be used to link some of these variables to preference-based utility values (e.g. HADs to EQ-5D). However, the functions currently available have not been validated on external data consequently it is recommended that the service user questionnaire also includes a measure which could be used to generate utilities.

9.4 Recommendations for psychological therapies

The NCA collects information from patients with anxiety or depression receiving psychological therapies in the community, and there is a mandatory field which could be used to differentiate between the subcomponents of this condition. The Service user questionnaire includes a measure to capture the strength of the therapeutic relationship between clients and their therapists,(88) and the retrospective service user questionnaire also includes information on measures such as the HADs, CORE-10 and Beck Anxiety Inventory (BAI). However, there is no measure which could be used to generate preference-based scores directly. Potential recommendations (PR) and areas for future research (FR) are provided below (Table 12) with the corresponding narrative provided in Appendix H.

The psychological therapies NCA data is currently being analysed under a separate research project (WP3), and the results of this research will inform additional recommendations for the fields in the audit (FR.20).

Table 12: Recommendations and associated future research for psychological therapies

PR.24	*Collect the EQ-5D in the service user questionnaire alongside clinical measures such as the PHQ-9 and HAD-7*
PR.25	*Collect the ReQOL in the service user questionnaire once it becomes available*
FR.19	*Assess the appropriateness of the EQ-5D and the ReQoL in patients receiving psychological therapies using the data from the NCA*
PR.26	*All therapists use a common set of measures (to be decided and ultimately synchronised with the measure adopted for use in NHS Outcome Framework)*
PR.27	*Include additional mandatory fields in the psychological therapies NCA*
FR.20	*Detailed analyses of fields currently collected in the NCA is currently being undertaken under a separate research project within this programme of work (WP.3).*

10. RESULTS FOR SCHIZOPHRENIA

10.1 Evidence of appropriatness of EQ-5D in schizophrenia (WP1.1)

An existing systematic review was updated (89) and a total of 10 studies were included in the new updated review. Of the ten included studies, two studies used the UK EQ-5D tariff, one used Finnish weights and the remainder were not detailed in Papaioannou et al.(90) Mean ages ranged from 28.9 (89) to 41.5 years,(91) and a fairly broad range of psychosis disorders were covered in the studies. The majority of measures used to compare and assess the EQ-5D were designed for use in mental health conditions, or to capture mental health symptoms.

The construct evidence (known groups) was good with two studies reporting that the EQ-5D detected differences in the expected direction for known groups, characterised by function (GAF), severity (PANNS) and condition (ICD).(90;92) However, the evidence for both responsiveness and convergent validity was mixed. For responsiveness, while there was some evidence that the EQ-5D was responsive to change, this evidence was limited to the PANSS positive subscale, the Groningen social disabilities schedule (GSDS) and the auditory hallucinations rating scale (AHRS), and no association was found when changes in the BPRS were small (<25%). Despite one study reporting moderate to large effect size for both symptom and function measures, the relationship between the EQ-5D and symptoms (function) was reported as poor in three (three) studies. In conclusion, there is sufficient evidence to raise doubts about the appropriateness of the EQ-5D in patients with schizophrenia (Table 13).

4.2 Evidence of alternative measures in schizophrenia (WP1.2)

Searches identified three reports of relevance to WP1.2.(93-95) All three recommended the same measures, namely the brief psychiatric rating scale (BPRS) and the positive and negative syndrome scale (PNASS). The EMA report states that these are reliable and validated measures, but does not provide evidence to support this statement.(95) The HoNOS and Friends and Family Test are recommended as suitable CROM and PREM respectively while the Warwick Edinburgh Mental Well Being Scale (WEMWBS) or the short version (s-WEMWBS) is currently recommended and both are being tested by the Care Pathways and Packages Project.(96) As mentioned earlier a new PROM, the ReQoL, is currently under development and will be suitable for use across the psychotic (which will include anxiety and depression) and non-psychotic (which will include schizophrenia) conditions. Once the measure is available and has been validated in people with schizophrenia, the ReQoL may become a candidate measure for inclusion in the NCA.

Table 13: Summary of evidence on measures for schizophrenia

Condition	N	Acceptability	Reliability	Construct		Responsiveness		Overall
				KGV	Convergent	Change over time	Ceiling Effect	
EQ-5D	11	NR	NR	Good	Mixed	Mixed	NR	Not appropriate
HoNOS (clinician-completed)		The recommendation is based on those in PBR [DH2013] and the psychometric properties of this measure have not been reviewed in the current report						
WEMWBS		The recommendation is based on those in PBR [DH2013] and the psychometric properties of this measure are currently under review elsewhere (96)						
ReQOL		This measure is currently in development and will be available in 2015						

N= number of studies used to inform conclusions, KGV: known group validity; NR, the existing review did not review this psychometric property.

10.3 Evidence for economic evaluations in schizophrenia (WP1.3)

Just one STA relating to schizophrenia was identified.(97) The evaluation compared pharmacological treatments for schizophrenia in adolescents (15-17 years). As treatment related adverse events such as substantial weight gain and somnolence are prevalent, and frequently lead to discontinuation of treatment, these were captured within the model framework. QALYs were obtained by assigning mean utility values to the discrete health states and due to lack of more suitable data in adolescents, EQ-5D data collected from adults were used. The results of the searches conducted to inform the model parameters suggest the volume of EQ-5D data in patients with schizophrenia is very limited and none were available in adolescents.

The mandatory fields in the schizophrenia NCA do not provide sufficient detail to model the individual treatment effects (maintenance, relapse, weight gain etc) as applied in the existing HTA cost-effectiveness evaluations in schizophrenia. Although remission (full or partial) is a mandatory field, it is believed these records are subjective clinical decisions. If this is the case, this evidence could be improved through the use of a clinical instrument with clearly defined criteria for remission. The rate of patients in remission could then be used to inform economic models comparing interventions or providers. Some of the optional fields in the clinical audit tool, such as the use of antipsychotic medications and history of medications in patients not in remission, could supply some of the additional evidence required to model the cost-effectiveness of different interventions and policies. There are currently no data collected in the schizophrenia NCA which could be used to inform the HRQoL associated with the condition or the interventions prescribed and the inclusion of

a variable which could be used to generate preference-based utilities would greatly enhance the dataset.

10.4 Recommendations for schizophrenia

Based on the evidence reviewed, the EQ-5D is not thought to be appropriate for patients with schizophrenia. It is not believed that there are data in the schizophrenia NCA which could be used to inform the HRQoL associated with the condition, either directly through a preference-based measure, or indirectly through an alternative measure. In addition, it is not believed that the other variables collected in the audit will suffice to compare providers or conduct robust economic evaluations. Potential recommendations (PR) and areas for future research (FR) are provided below (Table 14) with the corresponding narrative provided in Appendix I.

Table 14: Recommendations and associated future research for schizophrenia

PR.28	*Include both the HoNOS (a clinician-completed measure) and WEMWBS in the schizophrenia NCA*
PR.29	*Include the ReQOL in the NCA once available and validated in patients with schizophrenia*
FR.21	*Assess the psychometric properties of the ReQOL using the data collected in the NCA*
PR.30	*Increase the mandatory fields in the NCA to facilitate future economic evaluations*
FR.22	*Inspect the fields collected in the NCA with a view to making recommendations on the information required to compare providers and conduct economic evaluations*

11 RESULTS FOR DEMENTIA

11.1 Evidence of appropriateness of EQ-5D in dementia (WP1.1)

An existing review which appraised the evidence on the appropriateness of the EQ-5D in dementia (n=18) was used. Although some positive results were reported, in general, there was sufficient evidence to raise concerns relating to the appropriateness of EQ-5D in patients with dementia (Table 15). A ceiling effect was observed in three studies,(98-100) two studies reported the EQ-5D may not be acceptable for patients with severe dementia,(99;101) and two studies reported no relationship between self-reported EQ-5D scores and clinical measures.(98;102) Conversely two of three studies reported there was a relationship between proxy scores and clinical variables.(98;103) However, several issues with proxy scores were also described. Six studies reported no relationship between self-reported and proxy scores (even in patients with mild dementia).(98;100-102;104;105) Patients scored higher HRQoL than proxies in those that provided this information and the carers' responses were influenced by the level of dependency of the patient.(100;102;105) In addition, three studies reported no association between clinician and carer-proxy scores, with evidence suggesting that each may have a more accurate concept of particular attributes of HRQoL.(99;101;106)

Table 15: Summary of evidence on EQ-5D for dementia

Measure (N)	Acceptability	Reliability	Construct (KGV; Convergent)	Responsiveness (Change over time; Ceiling effects)	
Self-reported					
EQ-5D (17)	Good for mild dementia. Poorer for moderate to severe dementia.	-	Mixed; Poor	-	Evidence of ceiling effects
Proxy-rated					
EQ-5D (19)	Good	-	Some positive evidence, but methods have flaws; Fair.	-	Evidence of ceiling effects
Concerns about appropriateness.					

11.2 Alternative measures in dementia (WP1.2)

Searches identified two potentially relevant documents, neither of which makes specific recommendations on the selection of instruments to measure HRQoL, despite discussing the psychometric properties and appropriateness of different instruments (107;108).

In addition to the documents mentioned in the previous paragraph, evidence presented in two manuscripts known to the authors is worthy of consideration.(109;110) The use of a bolt-on for cognitive impairment has been suggested for EQ-5D,(288) but the proxy-rated EQ-5D+C was found to perform similarly to the EQ-5D and its use in isolation is not recommended. (286) The second source of evidence, an HTA report on two dementia specific HRQoL instruments, the DEMQOL and the DEMQOL-proxy (patient and proxy rated), provides early positive evidence of acceptability, validity and responsiveness of these instruments, but requires further validation in datasets incorporating a range of clinical indicators and dementia severity levels before they can be considered appriate for use in dementia.(287)

11.3 Evidence for economic evaluations in dementia (WP1.3)

11.3.1 *Cost-effectiveness modelling approach used in recent HTAs in dementia*

One MTA which assessed the clinical and cost-effectiveness of anti-dementia medication (acetylcholinesterase inhibitors, i.e. donezepil, galantamine and rivastigmine, and memantine) compared to each other and best supportive care for the treatment of Alzheimer's disease was reviewed.(111) The natural disease history was modelled by two multivariate regression time to event models (time to institutionalisation and time to death), which predict events based on age, cognition (measured by the Mini-mental State Examination (MMSE)) and functional ability (measured by Activities of Daily Living (ADL)). The model quality adjusted survival by assigning mean utility values to the discrete health states.

The TA on Alzheimer's disease(111) used survival analysis to model mortality and disease progression. The audit collects data of in-hospital death, as well as patient's age which was a covariate in the regression models used to predict mortality and time to institutionalisation in the economic model. However, the National Audit of Dementia (NAD) does not collect data on the two other covariates in both regression models, i.e. measures of cognition and physical ability. There are two fields on the casenote audit that provide some information on whether the patient was institutionalised before and/or after hospital admission. The only treatment data collected by the audit refers to the use of antipsychotic drugs, which are not anticipated to impact on disease progression, but may reduce behavioural and psychological symptoms.

Data collection within the NAD has the objective of allowing comparisons between hospitals in terms of standard of treatment and care provided to dementia patients. The focus of the audit is mostly on describing the treatment, care and support of these patients. Although this is valuable information,

the audit in its current format does not collect any variable that can be used to derive utility estimates, directly or indirectly (e.g. through a mapping function). To our best knowledge the collection of any PROMs within the NAD is not currently being considered.

11.4 Recommendations for dementia

Based on the evidence reviewed, the EQ-5D is not thought to be appropriate for patients with dementia. It is not believed that there are data in the NAD which could be used to inform the HRQoL associated with the condition, either directly through a preference-based measure, or indirectly through a surrogate. In addition, it is not believed that the other variables collected in the audit will suffice to conduct robust economic evaluations. PPR and areas for FR are provided below (Table 16) with the corresponding narrative provided in Appendix J.

Table 16: Recommendations and associated future research for dementia

PR.30	*Collect the* DEMQOL-U and the DEMQOL-U-proxy *in a service user questionnaire alongside clinical measures such as the MMSE and ADL.*
FR.23	*Assess the appropriateness of the DEMQOL-U and the DEMQOL-U-proxy in dementia patients using the data from the NCA*
PR.31	*Collect mandatory information on time and date of full-time institutionalisation, type of drug therapy administered, death and utility values.*

12 RESULTS FOR CARDIAC ARRHYTHMIA

12.1 Evidence of appropriateness of EQ-5D in cardiac arrhytmia (WP1.1)

Evidence of the appropriateness of the EQ-5D is presented jointly for all cardiovascular disease (CVD) conditions considered in the report. This was considered to be an appropriate way to present results, given i) the paucity of evidence available for each individual cardiovascular condition, and ii) the existence of some overlap of the study populations as defined by the cardiovascular conditions (e.g. acute coronary syndrome (ACS)) patients can be treated with coronary angioplasty or cardiac surgery, such as coronary artery bypass graft (CABG).

An existing review was updated,(112) and 12 studies were included in the current review.(113-124) The majority of studies applied the EQ-5D UK tariff to estimate utility weights. Five of the studies assessed the psychometric properties of the EQ-5D in ACS.(113;115;119;124;125) Five of the studies assessed the psychometric properties of the EQ-5D in heart failure.(116;117;120-122) One of the studies examined the acceptability of EQ-5D-5L in cardiac arrhythmia.(123) Four of the studies assessed the psychometric properties of the EQ-5D in a patient population that included more than one cardiovascular condition in WP1.1. There were three studies that included ACS, coronary angioplasty and cardiac surgery, and one study that included coronary angioplasty and cardiac surgery.(114)

Overall, the evidence base assessing the performance of the EQ-5D in cardiovascular conditions is mostly positive. The acceptability of EQ-5D was fair to good in the three studies that assessed this property. Evidence regarding reliability was very limited and indicated poor performance. Nevertheless, reliability was examined in one study alone and only at health dimension level. Construct validity (convergent) between EQ-5D was generally good, and more evident at the index than at the health dimension level. This was the psychometric property for which there was more robust evidence, given the amount of evidence available and its general concordance. The available evidence in terms of construct validity also reinforced that the EQ-5D can distinguish between subgroups of patients of varying HRQoL, namely between myocardial infarction (MI) and CABG patients. There was considerable positive evidence of known-group validity across groups defined by age, gender and educational level.The ability of EQ-5D to distinguish between groups of different disease severity was found to be good. Finally, evidence on responsiveness as change over time was poor across the three studies that assessed it, but it should be noted that the methodology used to examine itwas not the most adequate. It was concluded that the evidence on responsiveness is mostly uncertain, but studies so far suggest that EQ-5D may perform poorly in this characteristic.

Importantly, the studies that assessed responsiveness as change over time were all conducted in a single disease condition, heart failure, so there is also a question of whether the poor performance in this condition is likely to extend to other cardiovascular conditions. One study detected a ceiling effect, i.e. a tendency towards single level response, that was considerably more pronounced for individual health domains of EQ-5D (especially for self-care), than for the utility score.(114) Evidence of ceiling effects for EQ-5D utility scores was also found in a large European study.(125) This may translate into less discriminative ability of EQ-5D for patients at lower levels of disease severity.

Six studies included in the review assessed the psychometric properties of EQ-5D in heart failure, and provided evidence of construct (both convergent and known-group) validity, but poor responsiveness. For ACS and cardiac surgery there was some evidence of the majority of all reported psychometric properties (with the exception of responsiveness), taken from five and four studies respectively. EQ-5D performed well in terms of acceptability and validity, but showed poor reliability. In the five studies in coronary angioplasty, EQ-5D performed similarly to what was found for ACS and cardiac surgery in the same properties. The only evidence found for cardiac arrhythmias refered to acceptability of EQ-5D, which was considered good. Results by cardiovascular condition and psychometric property are summarised below in Table 17.

Table 17: Conclusions on evidence on psychometric properties of EQ-5D in CVD

			Construct	Responsiveness	
Condition	**Acceptability**	**Reliability**	**(KGV;Convergent**	**(Ceiling effects)**	**(Change over time)**
Cardiac arrhythmia	Good	NE	NE;NE	NE	NE
Heart failure	NE	NE	Good;Good	NE	Poor
Coronary angioplasty	Fair	NE	Good;Fair/Good	Potential ceiling effects	NE
Adult cardiac surgery	Fair/Good	Poor	Good;Fair/Good	Potential ceiling effects	NE
Acute coronary syndrome	Good	Poor	Good;Good	Potential ceiling effects	NE

KGV: known group validity; NE: no evidence

12.2 Routinely collected proxy measures in cardiac arrhythmias (WP1.2)

As the EQ-5D was found to be acceptable for CVD conditions, no additional searches for alternative measures were conducted.

12.3 Evidence for economic evaluations in cardiac arrhythmias (WP1.3)

Seven economic models from existing TAs were reviewed. Four of the TAs examined the clinical and cost-effectiveness of a pharmaceutical intervention in atrial fibrillation (AF),(126-129) while the other three assessed devices to manage cardiac arrhythmias.(130-132) Across the TAs, HRQoL in patients with cardiac arrhythmias has been assumed to depend mostly on underlying coronary disease and its severity, complications subsequent to surgical procedures, symptoms, and adverse effects of anticoagulant drugs.(126-132) Disability following stroke is another important aspect of HRQoL, especially in patients at increased risk of stroke, i.e. patients with atrial fibrillation (AF), with assigned utility weights varying by level of disability.(126-128)

Key clinical events in the existing models include death (all cause, cardiovascular and non-cardiovascular), short and longer term complications from the condition or intervention, and hospital readmissions. The National Audit of Cardiac Rhythm Management (NACRM) already collects some of these events, although their collection is not mandatory. Nevertheless, there are other key clinical events that would not be captured by this audit, such as readmissions to hospital due to clinical deterioration (e.g. hospitalisation due to heart failure worsening). Furthermore, the NACRM only collects data on antiarrhythmic drugs on patients with AF, with no other medication data being collected in the two datasets of the audit. These limitations could potentially be overcome by linking the NACRM data to datasets, such as some of the other NCA datasets (e.g. heart failure).

Patient related outcome measures are already collected in the NACRM, the EQ-5D and Atrial Fibrillation Effect on Quality Of Life questionnaire, although only for patients undergoing ablation procedures. The extension of the collection of PROM data to all patients in the NACRM (preferably the EQ-5D), could improve the ability of the audit to inform cost-effectiveness analysis of interventions. The value of collecting PROMs data in cardiac arrhythmias may be greater than in other conditions where it is possible to map clinical variables to a preference-based measure. In this particular case, mapping from clinical variables could be very challenging, given that these variables are usually disease specific and there are several underlying cardiovascular diseases that can be the cause of cardiac arrhythmias.

12.4 Recommendations for cardiac arrhythmias

In summary, no evidence was identified on the validity and responsiveness of the EQ-5D in cardiac arrhythmias, although the instrument was considered to have good acceptability (Section 12.1). Nevertheless, the validity of EQ-5D has been demonstrated in other related conditions, such as heart

failure, as well as more generally in the cardiovascular area. Furthermore, EQ-5D derived utility weights have been widely used in cost-effectiveness studies in cardiac arrhythmias, and NACRM already collects EQ-5D data in patients with atrial fibrillation who undergo atrial ablation procedures. In addition, and although many variables of importance are already collected in the audit, there are concerns about the completion rates of fields not included in the NACRM Minimum Data Standard and that not all relevant fields to perform robust economic evaluations are collected in the audit. The issues and corresponding PR and areas for FR are tabulated below (Table 18) with the corresponding narrative provided in Appendix K.

Table 18: Recommendations and associated future research for cardiac arrhythmias

PR.32	*Extend EQ-5D-5L collection so that it is conducted for both elements of the NACRM, i.e. the device and the ablation procedures dataset. implantation of cardiac devices (collected in the spreadsheet Device-dataset-5502n-14032014)*
FR.24	*Assess the psychometric properties of the EQ-5D-5L in patients with cardiac arrhythmias using data collected in the audit*
PR.33	*Collect mandatory information on the use of anticoagulant drugs and the occurrence of adverse events associated with these drugs, with special attention to bleeding and cerebrovascular events*
PR.34	*Depending on completion rates of Minimum Data Standard Dataset, consider making these fields mandatory*

An erratum has been posted online relating to the collection of the EQ-5D variable in the cardiac arrythmia audit (http://bit.ly/1V3aQjC).

13 RESULTS FOR HEART FAILURE

13.1 Evidence of appropriateness of EQ-5D in heart failure (WP1.1)

Six studies included in the updated review in CVD conditions (Section 12.1) assessed the psychometric properties of EQ-5D in heart failure, and provided positive evidence of construct (both convergent and known-group) validity, but poor responsiveness. Full details on the assessment of the appropriateness of EQ-5D in heart failure, and more generally in CVD, are presented in Appendix K.

13.2 Routinely collected proxy measures in heart failure (WP1.2)

As the EQ-5D was found to be acceptable for CVD conditions, no additional searches for alternative measures were conducted.

13.3 Evidence for economic evaluations in heart failure (WP1.3)

Of the three TAs relating to heart failure identified from the searches, two examined the clinical and cost-effectiveness of devices to manage cardiac arrhythmia in patients with heart failure.(130;131) A third TA examined the clinical and cost-effectiveness of pharmaceutical interventions for the treatment of chronic heart failure.(133) The HRQoL estimates applied to health states of the economic models in the three TAs were dependent on the severity of heart failure according to the New York Heart Association classification (NYHA) and corresponded mostly to EQ-5D utility scores.(130;131;133)

In the existing models, mortality (all cause, cardiovascular and non-cardiovascular) was a key clinical event. The National Heart Failure Audit (NHFA) collects mortality data, but will not allow distinction between cardiovascular and non-cardiovascular death, as cause of death is not collected within the audit. Furthermore, the NHFA will only provide data on death occurring at the hospital. Other key clinical events in the economic models, such as hospitalisation due to heart failure worsening could potentially be estimated by using readmission data collected in the audit. Transitions to stable heart failure states could also be modelled based on discharge data in the NHFA. Finally, the NHFA data on referrals can be used to inform further clinical events, namely cardiothoracic surgery and heart transplant, although its collection is not mandatory.

Importantly, the audit does not provide data on surgical complications and device related procedures, which are relevant for those patients requiring a device to manage cardiac arrhythmia or heart transplant. The identified gaps in mortality data could potentially be overcome by using

external data (e.g. Office of National Statistics). Similarly, data on events related to surgical complications and device related procedures could be obtained by linking the NHFA dataset to other national audit datasets, namely the Cardiac Rhythm Management datasets.

Patient related outcome measures are not currently collected in the NHFA. The inclusion of a preference-based HRQoL questionnaire (preferably the EQ-5D), could improve the ability of the NHFA to inform cost-effectiveness analysis of interventions. An alternative might be to map from a clinical variable currently collected in the NHFA, breathlessness (corresponding to NYHA classification for severity), to a preference-based measure, which would be compatible with the modelling approaches used in previous TAs. However, the NYHA is only collected at admission and readmission, and for mapping purposes it would be valuable to collect this measure at discharge too to capture potential benefits of interventions.

13.4 Recommendations for heart failure

In summary, the EQ-5D appears to be appropriate in patients with HF, and the current heart failure audit collects much of the information required to conduct economic evaluations. Nevertheless, the audit does not collect any HRQoL data, and death rates by cause are not collected. The issues and corresponding PR and areas for FR are tabulated below (Table 19) with the corresponding narrative provided in Appendix K.

Table 19: Recommendations and associated future research for heart failure

PR.35	*Collect the EQ-5D (EQ-5D-5L) in the NCA*
FR.25	*Assess the responsiveness of the EQ-5D-5L to changes in NYHA (already collected in the NCA)using data collected in the audit*
PR.36	*Collect mandatory information on cause of death*
FR.26	*Analyses of fields currently collected in the heart failure NCA is currently being undertaken under a separate research project within this programme of work (WP3)*

14 RESULTS FOR CORONARY ANGIOPLASTY

14.1 Evidence of appropriateness of EQ-5D in Coronary Angioplasty (WP1.1)

Five studies included in the updated review in CVD conditions (Section 12.1) assessed the psychometric properties of EQ-5D in coronary angioplasty, and provided positive evidence of acceptability, construct (both convergent and known-group) validity, but showed poor reliability. Full details on the assessment of the appropriateness of EQ-5D in coronary angioplasty, and more generally in cardiovascular disease, are presented in Appendix K.

14.2 Alternative measures in Coronary Angioplasty (WP1.2)

As the EQ-5D was found to be acceptable for CVD conditions, no additional searches for alternative measures were conducted.

14.3 Evidence for economic evaluations in Coronary Angioplasty (WP1.3)

Five TAs relating to coronary angiography were identified from the searches.(134-138) Three of the TAs examined the clinical and cost-effectiveness of pharmaceutical interventions for the treatment of a population that included patients who underwent revascularisation (namely percutaneous coronary intervention (PCI) and CABG.(134-136) Another TA examined clinical and cost-effectiveness of a pharmaceutical intervention, compared to heparin in addition to glycoproteins inhibitors for patients with ST elevation myocardial infarction (STEMI) intended for PCI.(137) Finally, TA 152 examined the clinical and cost-effectiveness of drug-eluting stents vs. bare metal stents was compared in patients with coronary heart disease.(138) All studies quality adjusted survival by assigning mean utility values to the separate health states, the majority of which were derived from mean EQ-5D estimates. In one model utilities were also weighted according to the level of subsequent disability in stroke health states.(134)

The National Audit of Percutaneous Coronary Interventional Procedures (NAPCI) collects data on all cause mortality that can be used to model survival in accordance with previous TAs. However, cause of death is not collected within the audit, and therefore it is not possible to distinguish between cardiovascular and non-cardiovascular mortality. Furthermore, the audit only collects data on death occurring at the hospital. The NAPCI audit collects data on PCI complicationsand revascularisation procedures (CABG and PCI) following initial PCI that can be used to inform the occurrence of further clinical events within the same hospital episode. There are other fields that also collect data related to complications, but these are not part of the minimum standard data. Importantly, the audit does

not collect data on all types of stroke or data related to severity of disability following cerebrovascular events.

Limitations on the collection of mortality data could be partially overcome by incorporating external data, namely by linkage to mortality registers, or by using estimates from the published literature. Linking the PCI audit to other datasets could not only allow the calculation of longer term mortality rates, but also the distribution of patients between cardiovascular and non-cardiovascular deaths. Even if mortality registers do not allow identifying cause of death, this can be ascertained through linkage to other audits or registers (e.g. Hospital Episode Statistics). A potential alternative approach to overcome gaps in mortality data might be to use external evidence on the expected rate of non-cardiovascular mortality. Clinical events occurring after hospital discharge or not recorded in the audit could also be sourced from the published literature or by linking the NAPCI data to other audits which collect data related to cardiovascular events (e.g. the Myocardial Ischaemia National Audit Project (MINAP) for ACS, and the National Adult Cardiac Surgery Audit (NACSA) for revascularisations with CABG).

The NAPCI does not currently collect patient related outcome measures. The inclusion of a preference-based HRQoL questionnaire (preferably the EQ-5D), could improve the ability of the audit to inform cost-effectiveness analysis of interventions. An alternative would be to apply values from the published literature to clinical events, as has been done in previous models. However, as mentioned before, the audit only collects data on those events that occur during the time spent in hospital following the initial PCI. Nevertheless, there is a very important limitation to this approach, as it may fail to capture the impact on HRQoL of the interventions, given that it may not be related to the occurrence of clinical events alone, but also to any beneficial effect on symptoms (e.g. relief of chest pain caused by myocardial ischaemia). Collection of a PROM would have a greater potential value to this particular audit, as it would allow HRQoL considerations to be related to events, but also to symptomatic differences that may be evident and are not fully accounted for by events. As it is, it may be difficult to demonstrate the full symptomatic benefits of intervention over just the hospitalisation period following coronary angioplasty, PROM collection should not be limited to this period. but also include longer follow-up intervals.

14.4 Recommendations for Coronary Angioplasty

In summary, the EQ-5D appears to be appropriate in patients undergoing coronary angioplasty, and the current NAPCI collects some of the information required to conduct economic evaluations. Nevertheless, the audit does not collect any HRQoL data, and could be improved by the inclusion of more fields and/or making their collection mandatory. The issues and corresponding PR and areas for FR are tabulated below (Table 20) with the corresponding narrative provided in Appendix K.

Table 20: Recommendations and associated future research for coronary angioplasty

PR.37	*Collect the EQ-5D (EQ-5D-5L) in the NCA prior to procedure, after procedure and at least one longer follow-up time point*
FR.27	*Assess the psychometric properties of the EQ-5D-5L using data collected in the audit*
PR.38	*Collect mandatory information on cause of death, type of stroke and severity of disability following stroke*
PR.39	*Make collection of data on drug therapy, number and type of stents and type of arterial access mandatory or at least part of the Minimum Data Standard dataset*

15 RESULTS FOR CARDIAC SURGERY

15.1 Evidence of appropriateness of EQ-5D in cardiac surgery (WP1.1)

Four studies included in the updated review in CVD conditions (Section 12.1) assessed the psychometric properties of EQ-5D in cardiac surgery, and provided positive evidence of acceptability, construct (both convergent and known-group) validity, but showed poor reliability. Furthermore, there was evidence of potential ceiling effects for the EQ-5D in this condition. Full details on the assessment of the appropriateness of EQ-5D in cardiac surgery, and more generally in CVD, are provided in Appendix K.

15.2 Routinely collected proxy measures in cardiac surgery (WP1.2)

As the EQ-5D was found to be acceptable for CVD conditions, no additional searches for alternative measures were conducted.

15.3 Evidence for economic evaluations in cardiac surgery (WP1.3)

The searches identified two TAs relating to cardiac surgery. Both TAs examined the clinical and cost-effectiveness of a pharmaceutical interventions in populations that included patients managed with PCI or CABG,(136) or have an initial coronary angiography for diagnostic purposes, and are then allocated to a primary treatment intervention with the majority of patients undergoing PCI.(137) None of these TAs assessed cardiac surgery procedures, and only include one type of procedure collected in the NACSA, namely CABG. Nevertheless, CABG is the most frequent procedure for which data is collected in the NACSA.(139) The two TAs quality adjusted survival by assigning mean EQ-5D utility values to the separate health states.

Similarly to other CVD audits, the data collected in the NACSA allows modelling all cause mortality, but does not allow distinguishing between cardiovascular and non-cardiovascular mortality. Furthermore, the collection of in-hospital death alone can limit the use of the audit to inform cost-effectiveness analysis. The NACSA also collects data that will inform the occurrence of several complications and further clinical events following initial surgery, although this is limited to the period of hospitalisation. Limitations on the collection of mortality data could be partially overcome by incorporating external data, namely by linkage to mortality registers, or by using estimates from the published literature. Linking the NACSA to other datasets could not only allow the determination of longer term mortality rates, but also the distribution of patients between cardiovascular and non-cardiovascular deaths. Even if mortality registers do not allow the identification of the cause of death, this can be ascertained through linkage to other audits or registers (e.g. Hospital Episode

Statistics), as described for coronary angioplasty in Section 14.3. Clinical events occurring after hospital discharge or not recorded in the audit could also be sourced from the published literature or by linking this NACSA to other CVD related audits. The considerations in this section apply mostly to CABG. Linkages to other audits and registers may be of relevance for less frequent cardiac surgery procedures, namely valve repair/replacement, aortic procedures and heart transplant, but identifying them would require further exploration of the clinical pathways following these procedures.

The NACSA audit does not currently collect patient related outcome measures. The inclusion of a preference-based HRQoL questionnaire (preferably the EQ-5D), could improve the ability of the audit to inform cost-effectiveness analysis of interventions. Similarly to the coronary angioplasty case, an alternative would be to apply values from the published literature to clinical events, as has been done in previous models. However, the audit does not collect all relevant clinical events, such as further ACS, and only collects data on those events that occur during the time spent in hospital following the initial PCI. As for coronary angioplasty, this approach may fail to capture the full impact on HRQoL attributable to the interventions, as HRQoL may not be related to the occurrence of clinical events alone, or to any beneficial effect on symptoms.

Although NYHA and Canadian Cardiovascular Society classification scores are collected in the NACSA, which capture some of the symptomatic dimension and can potentially be mapped into HRQoL measures, the collection is limited to stable patients in the pre-operative stage, and therefore not very useful in this context. The collection of a PROM would allow HRQoL considerations to be related to events, but also to symptomatic differences. As it is, it may be difficult to demonstrate the full symptomatic benefits of an intervention over just the hospitalisation period following cardiac surgery, PROM collection should not be limited to this period, but should also include longer follow-up intervals.

15.4 Recommendations for cardiac surgery

In summary, the EQ-5D appears to be appropriate in patients undergoing cardiac surgery, and the current NACSA collects some of the information required to conduct economic evaluations. Nevertheless, the audit does not collect any HRQoL data, and could be improved by including more fields and/or making some existing fields collection mandatory. The issues and corresponding PR and

areas for FR are tabulated below (Table 21) with the corresponding narrative provided in Appendix K.

Table 21: Recommendations and associated future research for cardiac surgery

PR.40	*Collect the EQ-5D (EQ-5D-5L) in the NCA prior to procedure, after procedure and at least at one longer follow-up time point.*
FR.28	*Assess the psychometric properties of the EQ-5D-5L using data collected in the audit*
PR.41	*Collect mandatory information on cause of death, type of stroke, severity of disability following stroke, and ACS and PCI on follow-up.*
PR.42	*Depending on completion rates of completeness assessment dataset, consider making these fields mandatory*

16 RESULTS FOR ACUTE CORONARY SYNDROME

16.1 Evidence of appropriateness of EQ-5D in acute coronary syndrome (WP1.1)

Five studies included in the updated review in CVD conditions (Section 12.1) assessed the psychometric properties of EQ-5D in acute coronary syndromes, and provided positive evidence of acceptability and construct (both convergent and known-group) validity, but showed poor reliability. Full details on the assessment of the appropriateness of EQ-5D in acute coronary syndrome, and more generally in CVD, are provided in Appendix K.

16.2 Alternative measures in acute coronary syndrome (WP1.2)

As the EQ-5D was found to be acceptable for CVD conditions, no additional searches for alternative measures were conducted.

16.3 Evidence for economic evaluations in acute coronary syndrome (WP1.3)

Five TAs relating to ACS were identified from the searches.(134-137;140) Four of the TAs examined the clinical and cost-effectiveness of paharmaceutical interventions for the treatment of people who have suffered an ACS.(134-137) Another TA examined the clinical and cost-effectiveness of statins for the prevention of coronary heart disease (including ACS).(140) All studies quality adjusted survival by assigning mean utility values to the separate health states, the majority of these values having been obtained via EQ-5D. One model weighted utilities according to levels of disability in the stroke health states. (134)

The data collected in MINAP could be used to model the different types of mortality (all cause, cardiovascular, and non-cardiovascular), but will only provide data on death occurring at the hospital. The MINAP also collects data on complications following ACS (bleeding, reinfarction), as well as reperfusion procedures (for STEMI patients) that can be used to inform the occurrence of further clinical events within the same hospital episode. Moreover, the audit does not provide data on the occurrence of stroke related events, and the level of disability resulting from these. Limitation on the collection of mortality data, could potentially be overcome by incorportating external published data or by linkage to other datasets (mortality registers, Hospital Episode Statistics, other audits) or longitudinal linkage across multiple MINAP entries (for further hospitalisations due to ACS). Another alternative is to use risk scores, such as the Global Registry of Acute Cardiac Events, to estimate cardiovascular mortality, as well as the risk of future MI events, in hospital and at 6 months.(141) Another limitation of the MINAP to inform cost-effectiveness studies, is that it provides incomplete coverage of other relevant clinical events that may occur during the initial

episode. This limitation could potentially be addressed by linking data to other audits, such as the NAPCI.

The MINAP does not currently collect patient-reported outcome measures. The inclusion of a preference-based HRQoL questionnaire (preferably the EQ-5D), could improve the ability of the audit to inform cost-effectiveness analysis of interventions. An alternative would be to apply values from the published literature to clinical events, as has been done in previous models. However, as it has been highlighted before, the MINAP does not collect all relevant data regarding clinical events, namely the occurrence of ischaemic and haemorrhagic strokes and level of disability following stroke. Even if the audit collected all relevant clinical events, applying utility weights to the clinical events alone may fail to capture the impact on HRQoL of the interventions if this extends beyond changing the frequency of the events and has an effect on symptoms too. Similarly to the PCI audit, collection of a PROM would allow HRQoL considerations to be related to events, but also to symptomatic differences that may be evident and are not fully accounted for by events. The issue regarding duration of follow-up period for the collection of PROMs in the NAPCI also applies here, as it may be difficult to demonstrate the full symptomatic benefits of interventions over just the hospitalisation period following ACS. Therefore, the collection of PROMs beyond the initial hospitalisation episode would be useful to ensure that any longer term symptomatic impact of treatments is captured.

16.4 Recommendations for acute coronary syndrome

In summary, the EQ-5D appears to be appropriate in ACS patients, and the current MINAP collects some of the information required to conduct economic evaluations. Nevertheless, the audit does not collect any HRQoL data, and could be improved by including more fields and/or making their collection mandatory. The issues and corresponding PR and areas for FR are tabulated below (Table 22) with the corresponding narrative provided in Appendix K.

Table 22: Recommendations and associated future research for acute coronary syndrome

PR.43	*Collect the EQ-5D (EQ-5D-5L) in the NCA at admission and at least one longer-term follow-up time point.*
FR.29	*Assess the psychometric properties of the EQ-5D-5L using data collected in the audit*
PR.44	*Collect mandatory information on occurrence and type of stroke, severity of disability following stroke.*

17. AGE AND COMORBIDITIES

The evidence available on age was extracted as reported in the reviews, or from the primary studies (in the case of the reanalysis, updates and the new review). With the exception of dementia (mean age range 61-81 years) and diabetes (mean age range 55-67 years), the subjects included in the primary studies covered a very broad range across all the conditions. With the exception of the mental health condition, no reference was made to the possibility of the EQ-5D being less responsive in older age groups. Consequently it is assumed that the results and conclusions drawn will generalise across a large age spectrum.

The presence of comorbidities in subjects was not available from either the reviews or from most of the primary studies reviewed, thus the reported exclusion criteria were used as an indicator of the possible presence of comorbidities. For example if the presence of comorbidities was not identified as an explicit exclusion criteria in the primary study, it was assumed that subjects with comorbidities were included in the primary studies. The only exclusion criteria identified which related to the presence of comorbidities were: additional specific mental health related conditions such as the presence of schizophrenia (in the published metal health review and in the diabetes studies),[Peasgood; Janssen] or the history and severity of cardiovascular conditions (in the updated cardiovascular review). Based on the absence of broader comorbidity related exclusion criteria, the broad age ranges covered, and the fact that the prevalence of comorbidities rises steeply by age, it is assumed that the results and conclusions drawn will generalise across subjects with comorbidities.

18. SUMMARY

The following section provides an overview of the results presented within the individual sections of the report. A summary of the evidence used to inform the conclusions for WP1.1 and WP1.2 is provided in Table 23 and 24. The report concludes with recommendations and suggested areas for future research. This section provides an overview only and it is recommended that the preceding sections and the corresponding Appendices are used for details on particular conditions.

18.1 Summary of evidence used to inform the conclusions for WP1.1 and WP1.2

IBD: A review of primary studies (n=2) provides evidence of acceptability, reliability, and known group/convergent validity for the EQ-5D in adults with IBD. However the evidence on the responsiveness of the EQ-5D is mixed with some ceiling effects and potential insensitivity to changes over time reported. While the EQ-5D is considered to be acceptable, additional validation is required particularly in patients with severe IBD and those undergoing surgical procedures. A review of evidence of PROMs for paediatrics provides evidence of acceptability, reliability and known group/convergent validity for the PedsQL™ (5 studies) in paediatrics with IBD. The PedsQL™ does not currently have an associated preference-base tariff, but it has both self-report and parent/carer versions and covers the full age spectrum for paediatrics (2-18 years). Additional preference-based measures are also recommended for use in paediatrics with IBD.

Epilepsy: A review of evidence of PROMs provides evidence of reliability and known group validity for the PedsQL™ (5 studies) in paediatrics with epilepsy, but the strength of the evidence supporting sensitivity to changes over time was less evident. While considered to be acceptable, additional validation is required to support the long term use of the PedsQL™ in this population. As with IBD, additional paediatric preference-based measures are recommended for use in paediatrics with epilepsy.

Diabetes: A reanalysis of an existing review (n=16 primary studies) provided evidence that the acceptability and reliability of the EQ-5D are good. There was some evidence of ceiling effects which may affect responsiveness in newly diagnosed diabetics and those without complications. Construct validity was generally good when compared to diabetes specific and generic quality of life measures, with a few exceptions, most notably in vision. Problems with the EQ-5D in vision have been noted elsewhere and addressed through the production of a vision "bolt-on" for the EQ-5D. It is recommended this is used alongside the EQ-5D. Paediatric measures were not reviewed due to time constraints and a low prevalence of diabetes in this population.

Bowel cancer: A reanalysis and update of an existing review (n=9 primary studies) provided mixed evidence regarding the construct validity and responsiveness of the EQ-5D. Acceptability, reliability

and ceiling effects were not reported by the original review authors. In some cases, known group validity appeared good against both cancer specific and generic measures, whilst in others it was unable to detect differences between groups where condition specific measures (EORTC QLQ-C30 or QLQ-CR38) could. One study reported convergent validity of the EQ-5D against TNM stages was low, as was the case for several other cancer specific and psychological symptom specific measures or subscales. However, some measures and subscales had moderate correlations. Responsiveness was also mixed, with the EQ-5D able to detect a change in health status over time in some cases, but not in others, where condition specific or symptom specific measures and subscales such as the HADS depression score, FACTC and EORTC measures did. It was concluded that the EQ-5D could not be recommended without further validation. A review of all alternative measures that could be used was not conducted, but it is recommended that the EQ-5D be used in conjunction with the EORTC-C30, which has a UK utility tariff, and the colorectal module EORTC QLQ-CR29 .

Head and neck cancer: One review was identified and an update search conducted. No primary research studies relating to the psychometric properties of the EQ-5D in head and neck cancer were identified. Searches identified seven published clinical or research guidelines relating to other measures, but none of these were based on up to date evidence. Given the limited evidence available, it is recommended that the EQ-5D is used alongside the EORTC QLQ-C30 and head and neck specific module, the QLQ-H&N35, in keeping with the recommendations for bowel cancer.

Psychological therapies: An existing review provided evidence from 20 primary studies relating to the EQ-5D in psychological therapies. Construct validity (both known group and convergent) was reported to be good, though data from one study showed that the EQ-5D correlated better with depression-specific measures and subscales than with anxiety-specific ones in people with anxiety. This suggests that additional research is required to confirm the appropriateness of the EQ-5D in patients with anxiety. Responsiveness was more mixed, but generally good, though notably one study in the elderly showed poor responsiveness. Better responsiveness of the EQ-5D was observed at the lower end of the HRQoL spectrum (e.g. severe depression) when compared to the SF-6D, which was more sensitive to changes at the top end of the spectrum. Overall, the EQ-5D was considered appropriate for use in anxiety and depression, though further validation work is required in anxiety. Searches were conducted to identify other measures. In keeping with The Royal College of Psychiatry, the GAD-7 and PHQ-9 measures are recommended for use alongside the EQ-5D. ReQoL, a measure currently in development by EEPRU for use in psychotic and non-psychotic mental health conditions is due to be available in July 2015, and could be considered for use once available.

Schizophrenia: An existing review was updated. A total of ten primary research studies were identified. Evidence was mixed. Construct validity by known group (defined by severity, diagnosis

subgroup and function) was good, but both construct validity by convergent methods and responsiveness by change over time were mixed. The EQ-5D was responsive to change over time in two studies, but the correlations between change scores in another study were only significant between the EQ-5D and an affect subscale (PANSS positive subscale), a social function scale (GSDS) and an auditory hallucination scale (AHRS). Small changes were not reflected in the EQ-5D scores. There are sufficient concerns with the EQ-5D in this population to prevent its recommendation. Three guidelines relating to other measures were identified and all recommended using the BPRS and the PANSS. Within the recently introduced payment by results initiative, the patient-reported measures used are WEMWBS and s-WEMWBS, and a clinician-reported measure, HoNOS, is also used. It is recommended that the same measures are used for the NCA. In this population, it may be useful to include a clinician-reported measure alongside patient-reported measures. ReQoL, could be considered as an alternative once available.

Dementia: An existing review provided evidence that there are some concerns relating to the use of the EQ-5D in dementia, including ceiling effects, and a lack of relationship between self-report and clinical measures. This review focussed some attention on the convergent validity between self-report and proxy-report, showing there was no relationship between self and proxy reports even in mild disease, and no association between carer-proxy and clinician scores, perhaps due to each having better insight in different attributes. Other measures were considered in two guidelines, and two reports known to the authors. It is recommended that two dementia-specific preference-based measures, the DEMQOL-U and the DEMQOL-U-proxy, should be collected in the audit alongside cognition and functional ability measures (see summary for WP1.3)

Cardiovascular conditions: An existing review (2010)(112) updated and a total of 12 primary studies were included in the update. As there was substantial overlap between the study populations, and a very limited amount of evidence for some of the individual CVD conditions the evidence is summarised collectively. Overall, the review provides evidence that the EQ-5D is adequate in CVD, being acceptable in the majority of studies and having good construct validity (known group and convergent). There was some evidence of ceiling effects (although this is unlikely to be observed in the hospitalised patients within the CVD audits), and there was very little evidence on its reliability. Additional evidence was required on its sensitivity to detecting small changes in HRQoL over time.

In summary, evidence from four new systematic reviews of primary studies (IBD adults, IBD paediatrics, epilepsy paediatrics, bowel cancer), four updates of existing systematic reviews (epilepsy adults, diabetes, bowel cancer, CVD) and two existing systematic reviews (psychological therpay) were used to identify if the EQ-5D was appropriate in the 13 health conditions. No

evidence was found for head and neck cancer. Six reviews of the literature, augmented with literature known to the authors were used to identify alternative or additional PROMs for patients with epilepsy (paediatrics), IBD (paediatrics), schizophrenia, cancer, dementia, patients receiving psychological therapies.

The psychometric properties of the EQ-5D were found to be adequate in ten of the 13 conditions. The exceptions were epilepsy (where the PedsQL™ was recommended), schizophrenia (where WEMWBS, to be replaced by ReQOL was recommended) and dementia (where DEMQOL-U was recommended). While the EQ-5D has been shown to be appropriate in the majority of the 13 conditions, the literature used to inform these conclusions was limited for several conditions and there was evidence of a potential ceiling effect or insensitivity to small changes in HRQoL.

Table 23: Summary of evidence supporting the psychometric properties of EQ-5D in all conditions

Condition	N	Acceptability	Reliability	Construct		Responsiveness		Overall
				KGV	Convergent	Change over time	Ceiling Effect	
IBD	2	Good	Good	Good	Good	Mixed	Mixed	Acceptable but not appropriate for paediatrics
Epilepsy	5	NE	NE	Good	Good	Mixed	Poor	All evidence in adults. Not appropriate for paediatrics
Diabetes	16	Good	Good	Good	Mixed	Good	Poor	Acceptable*
Bowel Cancer	9	NR	NR	Mixed	Poor	Mixed	NR	Acceptable
Head and Neck Cancer	0	NE	NE	NE	NE	NE	NE	No evidence
Psychological therapies	21	NR	NR	Good	Good	Mixed	NR	Acceptable
Schizophrenia	11	NR	NR	Good	Mixed	Mixed	NR	Not appropriate
Dementia	21	Mixed	NR	Mixed	Poor	NR	Poor	Not appropriate
Cardiac arrhythmia	1	Good	NE	NE	NE	NE	NE	Acceptable
Heart failure	6	NE	NE	Good	Good	Poor	NE	Acceptable
Coronary angioplasty	5	Fair	NE	Good	Fair/good	NE	Poor	Acceptable
Cardiac surgery	4	Fair/good	Poor	Good	Fair/good	NE	Poor	Acceptable
Acute coronary syndrome	5	Good	Poor	Good	~Good	NE	Poor	Acceptable

N= number of studies used to inform conclusions, KGV: known group validity; *Not appropriate for DM related vision problems, or neuropathy; NE, no evidence was identified; NR, the existing review did not review this psychometric property;

Table 24: Recommended measure when the EQ-5D is not appropriate

Condition	Alternative measure	N	Acceptability	Reliability	Construct		Responsiveness		Overall
					KGV	Convergent	Change over time	Ceiling Effect	
IBD paediatrics[a]	PedsQL	5	Good	Good	Good	No evidence	No evidence	No evidence	Acceptable
	PedsQL GI module		This measure is currently being validated and will be available shortly						
Epilepsy paediatrics	PedsQL	5	No evidence	Good	Good	No evidence	Unclear	No evidence	Acceptable
	PedsQL epilepsy module		This measure is currently in development and it is unclear when it will be available						
Diabetes (daily management)	DHP		The recommendation is based on those in PBR [DH2013] and the psychometric properties of this measure have not been reviewed in the current report						
Diabetes (vision)	EQ-5D vision bolt on		This measure requires additional validation in a large dataset						
Bowel Cancer	EORTC QLQ-C30 EORTC QLQ-CR38/29		The psychometric properties of these measures have not been reviewed in the current report						
Head and Neck Cancer	EORTC QLQ-C30 EORTC QLQ-H&N35		The psychometric properties of these measures have not been reviewed in the current report						
Psychological therapies	PHQ-9 GAD-7		Recommended by the Royal College of Psychiatrists Recommended by the Royal College of Psychiatrists						
	ReQOL		This measure is currently in development and will be available in 2015						
Schizophrenia	HoNOS (clinician-completed)		The recommendation is based on those in PBR [DH2013] and the psychometric properties of this measure have not been reviewed in the current report						
	WEMWBS		The recommendation is based on those in PBR [DH2013] and the psychometric properties of this measure are currently under review elsewhere [DH3013]						
	ReQOL		This measure is currently in development and will be available in 2015						
Dementia	DEMQoL-U		These measures (patient and proxy rated) provided early positive evidence of acceptability, validity and responsiveness, but requires further validation in datasets incorporating a range of clinical indicators and dementia severity levels.						

N= number of studies used to inform conclusions, KGV: known group validity;

[a] consider the PedsQL GI module as an adjunct to the core measure

18.2 Summary of evidence required for use in economic evaluations (WP1.3)

IBD: The EQ-5D is currently collected in the audit, but as it is not collected at the same time as other key variables used in the economics (for example, surgery or flares in symptoms), its usefulness in comparing interventions is limited. It may be possible to use a clinical variable (for example the Crohn's disease activity index (CDAI) in patients with CD) and an existing relationship between the CDAI and EQ-5D to enable the NCA data to be used in economic evaluations. Despite the issue with the timing of collections, the EQ-5D would be useful when comparing providers and if the timings of data collection could be synchronised with the clinical data, then it could be used in standard economic evaluations. While the audit collects much of the information required to conduct economic evaluations, for example the aggregate numbers of surgeries and surgical complications could be used to compare providers, it is not clear if there is sufficient evidence to adjust for case-mix. There are also areas where additional evidence, if mandatory, would be beneficial for future economic evaluations. These include details of pharmaceutical interventions and associated response and relapse data collected at the same time as a clinical variable such as the Crohn's disease activity index, surgical rates including type of intervention, success rate and associated complications.

Epilepsy: The existing patient questionnaire collects PREMs rather than PROMs and there is no existing variable within the current audit which could be used to map to a preference-based measure for use in economic evaluations. While the PREMs could be used to compare providers, their use in economic evaluations is limited. The review of existing economic evaluations identified that there was no suitable preference-based evidence in the literature for paediatrics with epilepsy. It is thought that the current audit contains insufficient detailed evidence on seizure frequency and severity, pharmaceutical interventions (and associated response, withdrawal rates and adverse events), surgical interventions and death rates to conduct formal economic evaluations with these data.

Diabetes: The existing audit does not include a patient questionnaire. A PREM focussed questionnaire for patients receiving secondary care is currently being piloted, although this is not believed to cover patients treated in primary care. The evidence collected in this questionnaire will be useful when comparing providers. There is a relatively large evidence base on preference-based data in patients with diabetes which could be used to inform formal economic models. However, there are gaps in this evidence where data collected in the audit would be beneficial. In particular, for patients with diabetes related vision conditions and to capture the HRQoL associated with hypoglycaemic events. The audit collects much of the evidence required to conduct formal

economic evaluations and if the inpatient data could be linked to GP records this would expand the evidence available considerably. There would remain some issues relating to the timing of the data collection, but it is believed that these data could be used to inform formal economic evaluations.

Bowel cancer: The bowel cancer audit does not include a patient questionnaire thus PROMs are not currently collected and existing literature on preference-based data which could be used to inform formal economic evaluations is sparse. The audit collects some of the information required to compare providers and economic evaluations such as survival, progression and staging of disease (Dukes' stage). However, key variables such as toxicity due to chemotherapy and adverse effects of surgery and radiotherapy are not currently mandatory fields.

Head and neck cancer: The head and neck cancer audit does not include a patient questionnaire thus PROMs are not currently collected and there does not appear to be an alternative field which could be used to predict the required preference-based utility values. Assuming there is a relatively high completion rate, it is though that this audit collects much of the information required to derive survival curves for progression and recurrence of the disease but again key information may be missing such as condition severity (measured using information of lesions and required to case-mix when comparing providers), current pharmaceutical interventions (type of intervention, concomitant medications, remission rates, relapse rates, adverse events) and surgical rates and complications.

Psychological therapies: Although the audit for patients receiving psychological therapies does not collect PROMs, there is a service user questionnaire which could potentially be amended to include a PROM. Two measures (the Beck's depression index, and the Beck's anxiety index) are also collected in the audit and it is possible that this evidence could be used to predict preference-based utility data using existing published relationships. There also appears to be an 'outcome measure' within the retrospective audit but it is unclear what this measure is hence it is not possible to determine its usefulness. Relapse rates are high for this condition and compliance to therapy can be problematic. Together with severity of the condition, these are key variables within economic evaluations but it is not clear if there are currently any mandatory fields within the audit relating to these. This audit is currently being used as a case-study in an associated project (WP3), and the results of this project will provide an indication of what can be achieved with the data collected.

Schizophrenia: Although the audit includes a service user and carer questionnaire, these do not currently include a PROM, concentrating of experience of and satisfaction with the health services provided. These data will be useful when comparing providers, but cannot be used to inform economic evaluations. It is not believed that the mandatory fields in this audit provide sufficient detail to model individual treatment effects (maintenance, relapse, compliance, weight gain etc), but

some of the optional fields could provide some evidence on antipsychotic medications and history of medications in patients not in remission. While there is a mandatory field relating to 'relapse' it is believed this is a subjective clinical decision, thus it may not be possible to use this in economic models.

Dementia: In its latest round, the dementia audit does not include a service user or carer questionnaire and thus no PROMs are currently collected. The information collected in the audit would enable comparison of providers (i.e. hospitals) in terms of the standard of treatment and care provided to patients with dementia, and to compare the performance of the individual hospital over time. However, it is not clear if the data could be case-mixed using variables such as cognition and physical ability. In addition to evidence on HRQoL, to conduct formal economic evaluations, the audit would require additional detailed mandatory information such as dementia diagnosis, MMSE score, ADL score, type of pharmaceutical therapy administered and death rates.

Cardiac arrhythmia: The NACRM currently collects the EQ-5D and the condition-specific AFTEC in a patient questionnaire administered pre and post (6 month and 12 month) the ablation procedure. However, only patients undergoing ablation procedures complete the questionnaire and if this could be extended to all patients within the audit, this would increase the scope of the audit data in relation to performing economic evaluations and comparing providers. Although there is currently insufficient information in the mandatory fields to conduct formal economic evaluations, the data standard subset has additional information that could be used, subject to completion levels. In particular, the following information would ideally be required for informing economic models: normal sinus rhythm, permanent AF with uncontrolled symptoms, permanent AF with controlled symptoms, and death rates, type of intervention (CRT pacemaker, CRT defibrillator, dual-chamber or single chamber pace-makers, implantable cardioverter defibrillators) and associated success/complication rates, cardiac resynchronisation therapy, anti-coagulant drugs and thromboembolic, ischaemic and bleeding events.

Heart Failure: Although no PROMs are currently collected in the NHFA, it may be possible to utilise the NYHA breathless severity data to obtain proxy preference-based utility scores to generate QALYs in economic evaluations. However, the NHFA data are only collected on admission and re-admission. To inform the benefits of interventions, they would also need to be collected post intervention and on discharge from hospital. In addition, the collection of EQ-5D-5L directly within the audit would capture the benefits of interventions and procedures directly thus reduce the uncertainty inherent within mapping functions. Excluding HRQoL information, the current NHFA collects much of the information required to conduct formal economic evaluations and to compare

providers, and it is possible that the gaps identified (mortality and surgical complications) may be available in external datasets if these could be linked in some way.

Coronary angioplasty: Although the NAPCI does not collect PROMs, due to the discrete nature of the health states in the typical clinical pathway, it would be possible to utilise evidence in the literature to populate HRQoL values in economic evaluations. However, the inclusion of a PROM (preferably the EQ-5D-5L) within the audit would enable direct comparison of providers and interventions using the audit data. In addition, depending on the timing of collection, EQ-5D-5L collected via the audit, could provide useful information on the longer-term effects (for example 6 month and 12 month post discharge) on HRQoL associated with reductions in symptoms, rather than the immediate direct effect of specific interventions and procedures (i.e. during hospitalisation). Excluding HRQoL information, the NAPCI does collect much of the information required to conduct formal economic evaluations and to compare providers. Again it may be possible to use external datasets to supplement gaps in the evidence collected (e.g. mortality and surgical complication rates) for economic evaluations, but this would not be particularly informative when comparing providers.

Cardiac surgery: The NACSA audit does not collect PROMs, and as discussed for the NAPCI, while it may be possible to utilise evidence from the literature when conducting economic evaluations of interventions, this form of information is not particularly informative when comparing providers and it is recommended that the EQ-5D-5L is collected within the audit with follow-up data to capture the longer-term HRQoL benefits of interventions. Excluding the HRQoL information, the information collected within the audit would suffice to compare providers and would provide a substantial amount of the evidence required to conduct formal economic evaluations of interventions (assuming a relatively high completion rate for all fields). The exceptions are again the mortality information, surgical complications and longer term information on subsequent events. The latter may be available from external datasets if these could be linked in some way.

Acute coronary syndrome: The MINAP does not collect PROMs and as discussed for the PCI audit, while it may be possible to utilise evidence from the literature when conducting economic evaluations of interventions, this form of information is not particularly informative when comparing providers and it is recommended that the EQ-5D-5L is collected within the audit with follow-up data to capture the longer-term HRQoL benefits of interventions. Excluding the HRQoL information, the information collected within the audit would provide a considerable amount of the information required to model the cost-effectiveness of interventions and policies in ACS assuming relatively high completion levels.

In summary, while the evidence collected in the individual audits will allow comparison of providers in many cases, it is clear that the mandatory fields in most of the audits will not provide sufficient detailed information to perform formal economic evaluations. The main omission is the lack of PROMs which limits the flexibility of the data in terms of comparing either providers or interventions used in routing clinical practice. However, many of the audits contain optional fields which would be useful for economic evaluations and enforcing the collection of key variables is recommended in many of the audits. A recurrent issue relates to the level of detail collected and the timing of the variables collected. To be useful for economic evaluations, many of the variables used have to be synchronised in terms of timing of collection, and many need to be collected over periods of time to assess progression or relapse etc. An additional key issue which arises throughout many of the reviews is the collection of information relating to side effects of pharmaceutical interventions and adverse events associated with surgical procedures and radiotherapy. The audits could provide valuable information on these rates, and the effects they have on patients' HRQoL, when used and performed in routine clinical practice.

18.3 Summary of recommendations and associated areas for future research

The reviews undertaken in this project have produced a relatively large number of condition specific and generic recommendations. Many of the recommendations would require additional supporting research and the majority of the research could be conducted in due course using the data collected in the corresponding NCAs. The recommendations and areas for future research relate to the development of patient questionnaires in the NCAs that do not currently collect information from service users and minimum mandatory fields required to perform economic evaluations using the NCA data.

Patient questionnaires: The majority of the NCAs do not currently include a patient questionnaire (*bowel cancer, head and neck cancer, dementia, heart failure, coronary angioplasty, cardiac surgery, acute coronary syndrome*). While it is possible to source the HRQoL evidence required to inform preference-based utility values in economic evaluations from the literature for some of the conditions, it is recommended that priority is given to developing patient questionnaires for the bowel cancer and head and neck cancer audits, as the evidence in the literature in these conditions is very sparse. It is also recommended that the patient questionnaire used in the diabetes secondary care audit is adapted and modified for use in the diabetes primary care audits to collect HRQoL evidence relating to hypoglycaemic events (and fear of these events), which are not currently available in the literature. Finally, while the EQ-5D is collected in the cardiac arrhythmias audit

(currently collected in patients undergoing ablation procedures only), it is recommended that this is extended to include all patients.

There are just three of the 13 conditions where the EQ-5D is not recommended for use:

- Schizophrenia: where the WEMWBS is recommended, to be replaced by the ReQOL in due course
- Dementia: where the DEMQOL –U is recommended
- Epilepsy: where the PedsQL™ is recommended due to the candidate population being paediatrics.

Additional research would be required in these conditions to enable these data to be used to generate QALYs in economic models. However, these data would still provide useful information for comparisons across providers.

PROMs: With the exception of epilepsy which is a paediatric population only, the EQ-5D-5L is recommended to compare providers or conduct economic evaluations, noting the following points:

- include the new five level version of the EQ-5D in patient questionnaires to inform preference-based utility values for economic models and to compare providers (*IBD, diabetes (with vision bolt-on), bowel cancer, head and neck cancer, psychological therapies, all cardiovascular conditions)*
- include additional condition specific PROMs in addition to the EQ-5D (*PRO2/3 in IBD, EORTC QLQ-C30 in bowel and head and neck cancer (with corresponding condition specific modules), PHQ-9/GAD-7 and ReQOL in psychological therapies)*
- for paediatrics, include the PedsQL™ and age-related generic preference-based measures (CHU-9D, HUI2 and EQ-5D-Y) in the patient questionnaires to inform preference-based utility values for economic models and to compare providers *(IBD, epilepsy and possibly diabetes depending on prevalence of paediatrics)*

Mandatory fields to inform economic evaluations

Condition Severity: Many of the economic models require some assessment of condition severity either to stratify patients or to link to changes in HRQoL. In many of the conditions clinical severity would also be a useful measure to case-mix patients when comparing across providers for example when assessing surgical success rates or mortality. Recommendations include:

- Collect CDAI or CAI in IBD, Dukes' stage in bowel cancer, mucositis grade and information on lesions in head and neck cancer
- Collect severity and frequency of seizures in epilepsy
- Collect BCVA or visual acuity scores in diabetes
- Collect MMSE and ADL in dementia

Interventions: It is not clear if the details of current pharmaceutical interventions are collected in many of the audits and information on the type and dose of intervention would be required to compare the cost-effectiveness of these interventions. In some of the conditions the known side-effects of the pharmaceutical interventions are particularly potent and can have a considerable detrimental effect on HRQoL. Collecting this information on patients treated in routine clinical practice would add considerably to the current evidence base in terms of quantification of the full effects of these interventions. Similarly it is not clear if all the information that would be required to compare provider success rates relating to surgical interventions and radiotherapy is collected as mandatory information in all the audits. Specific and generic recommendations relating to which information should be collected include:

- current pharmaceutical therapy *(IBD, epilepsy, PCI, cardiac arrhythmia)*
- response and or relapse rates for pharmaceutical treatments *(generic)*
- adverse events associated with pharmaceutical interventions (*generic but in particular, biologics in IBD, chemotherapy in bowel cancer and head and neck cancer, antipsychotic drugs in schizophrenia, anti-arrhythmic and anticoagulant drugs in cardiac arrhythmias)*
- surgical success rates and associated complications *(IBD, bowel cancer, head and neck cancer, CVD surgical procedures)*
- radiotherapy success rates and adverse events *(bowel cancer, head and neck cancer)*
- procedure details: *Number and type of stents; type of vascular access (coronary angioplasty)*
- cardiac event rates on follow-up *(cardiac surgery)*

Clinical events: All the conditions reviewed are chronic /progressive conditions and, in many, patients are at risk of condition related clinical events. Event rates in the existing economic evaluations are often derived using risk functions based on clinical measures such as blood pressure, lipid levels and weight. While it is recommended that the audits are checked to ensure that the appropriate clinical markers are collected as mandatory fields, additional information in the

following areas would be informative both when comparing providers and when conducting economic evaluations:

- cardiovascular and cerebrovascular events *(occurrence, type and severity of disability following event in diabetes, coronary angioplasty, cardiac surgery, ACS)*
- cause of death *(IBD, epilepsy, diabetes, bowel cancer, head and neck cancer, heart failure, coronary angioplasty, cardiac surgery)*

Future research

Many of the suggestions for future research are generic across the conditions:

- assess the psychometric properties of the recommended PROMs using the data collected in the respective audits (*all audits*)
- synchronise the timing of collection of clinical and HRQoL evidence to enable the HRQoL data to be used in economic evaluations (*all audits*)
- conduct a detailed review of data collected in the current audits with a view to informing which variables should be collected as mandatory evidence to inform future economic evaluations (*all audits*).

Individual condition specific recommendations and areas for future research include:

- Consider collaborating with the developer of the PedsQL^{TM} to develop an associated paediatric preference-based measure *(IBD, epilepsy)*
- Conduct mapping functions between clinical variables/PROMs and preference-based measures to enable the data to be used in economic evaluations (*IBD, diabetes, bowel cancer)*
- Generate new preference-based weights for EQ-5D vision bolt-on (*diabetes*)
- Explore the effects of hypoglycaemic events on HRQoL (*diabetes*)
- Ensure all therapists use a common set of measures, which match those adopted in the NHS Outcomes framework (*psychological therapies*)

Limitations

Limitations relating to WP1.1

Whilst WP1.1 provides a thorough overview of the literature relating to the psychometric properties of the EQ-5D in the 13 NCA conditions, there are a number of limitations which should be borne in mind.

1) The review methodology was a review of reviews. There are weaknesses inherent in such reviews, which apply:

- A review of reviews is always subject to greater error as data has been extracted twice: once by original review authors, then again for the review of reviews.
- Review of reviews depend on the quality of the study design, and on the accuracy in the implementation and reporting of the original review team
- Not all information needed to assess the quality of searches and data extraction is always reported.

2) Other limitations specific to this review of reviews include:

- Interpretations and categorisations of psychometric properties vary across the literature, and in some cases the definitions may not always have matched the inclusion criteria and definitions of WP1.1. Without reanalysing every review, it was not possible to always omit the data that did not match our definitions.
- In some cases, it was not possible to ascertain whether the conclusions of the authors were supported by the evidence, due to a lack of primary data being presented in the review. Where this was particularly problematic, data was extracted from the primary studies. However, this was only done for severe cases of lack of transparency, and some lack of clarity may persist in some cases.
- Some reviews analysed data in a different manner than was planned for WP1.1, and sometimes their methods were not transparent.
- A combination of the above three points (definitions of psychometric; lack of primary data presentation; differences in analysis methods), in combination with inevitable subjectivity in synthesis may have resulted in different conclusions having been drawn by the review authors than we would have drawn ourselves. Where such issues were considered very problematic, however, a re-analysis of the review data, sometimes including primary data extractions, was conducted.

- Each review conducted searches differently. Whilst the majority were judged to be adequate, there was not a consistent approach across reviews, and some searches may have been more comprehensive than others.
- Not all reviews performed quality assessment of primary studies, and we did not include this formally in WP1.1 either. A consensus on how to quality assesses studies in psychometric analyses has not yet been achieved. Some methods simply look to the number of missing data points, and the match of the patient characteristics with the population of interest. Both of these types of quality have been considered in our synthesis, but a remarkable lack of information relating to missing data points (many primary studies reported completers only) made a meaningful assessment of quality of the included studies impossible both within our own review, and within reviews we have reported. Even where reviews did perform quality assessment, it is rarely clear how this was integrated into review conclusions.

3) Other limitations

The paediatric reviews in epilepsy and IBD were unfortunately not comprehensive due to time constraints. The search terms included known paediatric preference based measures and as such, not all evidence will have been identified relating to other condition specific or generic measures that are not preference based.

Limitations relating to WP1.2

A full review, or a review or reviews, of all generic or conditions specific measures was not possible for WP1.2 within the time available. Searches were instead limited to searching for recommendations relating to PROMS in a few key sources for each condition. In most cases, these sources were not evidence based, and the rationale for the choice of one measure over another not always clear. As a result, we have, in large part, relied on the expert-knowledge of ongoing research and current DH programmes to inform the recommendations for WP1.2 in this report. Whilst this is not evidence-based, it is arguably a more suitable and integrated approach in the given circumstances, producing prescient recommendations that will be relevant for longer.

Limitations relating to WP1.3

Due to time constraints, an exhaustive strategy was not used to identify ongoing or scheduled research exploring the use of PREMs or PROMs in the different NCAs. In many cases, the contacts at the individual Audits were asked if they knew of any ongoing research in these areas. If required, it is possible to go back to this question to develop a more exhaustive strategy to identify any planned or ongoing research in more detail.

In addition, the recommendations for mandatory fields within the individual audits are based on the questionnaires used, rather than an in-depth review of the evidence collected. Consequently the recommendations for additional fields are indicative only.

APPENDIX A - Methods see link - Appendix A - Methods

APPENDIX B - Searches see link - Appendix B - Searches

APPENDIX C: INFLAMMATORY BOWEL DISEASE (see separate document)

APPENDIX D: EPILEPSY (see separate document)

APPENDIX E: DIABETES (see separate document)

APPENDIX F: BOWEL CANCER (see separate document)

APPENDIX G: HEAD AND NECK CANCER (see separate document)

APPENDIX H: PSYCHOLOGICAL THERAPIES (see separate document)

APPENDIX I: SCHIZOPHRENIA (see separate document)

APPENDIX J: DEMENTIA (see separate document)

APPENDIX K: CARDIOVASCULAR DISEASE (see separate document)

APPENDIX L: AGE AND COMORBIDITIES

Additional detail on information relating to the ages and comorbidities described in the existing reviews.

Table L1: Information on age and comorbidities as reported in the reviews[a]

Condition[b]	Age in years	Comorbidities	Comments
IBD New review	Mean age ~42 (SD ~12)	No exclusions were reported in the primary studies	Given the relatively large variation in age, and the fact that specific exclusions relating to comorbidities were not reported, it is likely that subjects with comorbidities were included. Consequently the results should generalise across a broad range of ages and subjects with comorbidities.
Diabetes Reanalysis of Janssen et al	Mean age range: 55-67	4/16 studies did not report any exclusions in their recruitment criteria 1/16 selected only diabetics with high risk of vascular disease 11/16 studies excluded some of the following: Psychological illness (7); cardiac disorders (1; hepatic disease (1; pregnancy (1; severe uncontrolled hypertension (1); systemic corticosteroid treatment (1); recreational drug/alcohol abuse (2); any serious, unstable/acute medical condition (4); neurological disorders (4); other pain condition (4); skin condition (3); amputation (1); macular degeneration &/or glaucoma (1); hearing difficulty (1); HIV/AIDS (1)	Given the mean age ranges, comorbidities are very likely in the four studies that did not report any exclusion criteria. Of studies that were selective, one actively recruited those with diabetes and high risk of vascular disease, and did not list any other exclusion criteria. Of the remaining 11, there were exclusions of some comorbidities. Mostly, this occurred in only a few studies. However, psychological illness has been excluded from seven studies, usually for practical reasons. As such, the evidence base may not be generalizable to patients with serious psychological conditions.
Mental health Peasgood et al	Mean age range: 40-74 Minimum reported:17 Maximum: not reported	6/31 of the studies reported specific mental-health related exclusions such as history of mania, schizophrenia or suicide intention. 1/31 of the studies excluded anyone with a 'comorbid condition'[Aydemir et al, 2009]	One study (mean age 74.1) concluded the EQ-5D was less responsive in older patients as the EQ-5D was less responsive than the BDI-II. However, several of the studies recruited subjects representative of the general population, thus the

		No other exclusions were reported for the balance (24/31) of studies. 1 study reported 59% in primary care have ≥ 1 comorbidity[Sobocki 2007]	results should generalise across a broad range of ages and subjects with comorbidities.
Dementia Hounsome et al	Mean age range: 61-81	No exclusions were reported in Hounsome	Given the mean age range, it is likely that there are subjects with comorbidities in at least some of the studies and the results should generalise across.
Schizophrenia Pappaiannou et al	Mean age range: 29-41 Minimum reported:18 Maximum reported: 80	No exclusions were reported in Pappaiannou	Given the age range, it is likely that there are subjects with comorbidities in at least some of the studies and the results should generalise across.
Bowel cancer Longworth et al	Not reported in Longworth	No exclusions were reported in Longworth	No reporting or discussion of co-morbidities or age for review or from individual studies.
Cardiovascular disease (updated review)	Mean age range: 55-69 Minimum reported:17 Maximum reported: 88 3/13 studies excluded older subjects (>60 n=1 study, >80 n=2 studies) 7/13 studies excluded younger subjects (<17 n=1 study, <18 n=2 studies, <30 n=2 studies, <35 n = 2 studies)	Exclusions included: a) Age (see adjacent cell) b) Cardiovascular condition: duration of time since previous cardio event or intervention, severity of cardiovascular condition (i,.e. New York heart failure class IV)	Given the reported age ranges, and exclusion criteria, it is likely that there are subjects with comorbidities in at least some of the studies. Consequently the results should generalise across a broad range of ages and subjects with comorbidities.

[a] The data were extracted as reported in the reviews, and this evidence was not reported for all the individual studies. [b] Epilepsy: NCA for children hence presence of comorbidities is not applicable for the vast majority; Head and neck cancer: no studies were identified in the searches.

Reference List

(1) Longworth L, Singh J, Brazier J. An evaluation of the performance of EQ-5D: a review of reviews of psychometric properties. Unpublished report to EuroQol Group, editor. 2014. Ref Type: Personal Communication

(2) Wille N, Badia X, Bonsel G, Burstrom K, Cavrini G, Devlin N, et al. Development of the EQ-5D-Y: a child-friendly version of the EQ-5D. Qual Life Res 2010 Aug;19(6):875-86.

(3) Stevens K. Working with children to develop dimensions for a new preference-based, generic, pediatric, health-related quality of life measure. Qualitative Health Research 2010;20:340-51.

(4) Feeny D, Furlong W, Barr RD. Multiattribute approach to the assessment of health-related quality of life: Health Utilities Index. Med Pediatr Oncol 1998;Suppl 1:54-9.

(5) Varni JW, Seid M, Rode CA. The PedsQL: measurement model for the pediatric quality of life inventory. Med Care 1999 Feb;37(2):126-39.

(6) Konig HH, Ulshofer A, Gregor M, von TC, Reinshagen M, Adler G, et al. Validation of the EuroQol questionnaire in patients with inflammatory bowel disease. Eur J Gastroenterol Hepatol 2002 Nov;14(11):1205-15.

(7) Stark RG, Reitmeir P, Leidl R, Konig HH. Validity, reliability, and responsiveness of the EQ-5D in inflammatory bowel disease in Germany. Inflamm Bowel Dis 2010 Jan;16(1):42-51.

(8) Abdovic S, Mocic PA, Milosevic M, Persic M, Senecic-Cala I, Kolacek S. The IMPACT-III (HR) questionnaire: a valid measure of health-related quality of life in Croatian children with inflammatory bowel disease. J Crohns Colitis 2013 Dec 1;7(11):908-15.

(9) Duffy CM, Wells GA, Russell AS, Haraoui B. Quality of life issues in pediatric immune-mediated inflammatory disease. J Rheumatol 201138(SUPPL. 88):20-25. Available from: URL: http://jrheum.org/content/supplements/88/20.full.pdf+html

(10) Marcus SB, Strople JA, Neighbors K, Weissberg-Benchell J, Nelson SP, Limbers C, et al. Fatigue and health-related quality of life in pediatric inflammatory bowel disease. Clin Gastroenterol Hepatol 2009 May;7(5):554-61.

(11) Perrin JM, Kuhlthau K, Chughtai A, Romm D, Kirschner BS, Ferry GD, et al. Measuring quality of life in pediatric patients with inflammatory bowel disease: psychometric and clinical characteristics. J Pediatr Gastroenterol Nutr 2008 Feb;46(2):164-71.

(12) Upton P, Eiser C, Cheung I, Hutchings HA, Jenney M, Maddocks A, et al. Measurement properties of the UK-English version of the Pediatric Quality of Life Inventory 4.0 (PedsQL) generic core scales. Health Qual Life Outcomes 2005;3:22.

(13) Ogden CA, Akobeng AK, Abbott J, Aggett P, Sood MR, Thomas AG. Validation of an instrument to measure quality of life in British children with inflammatory bowel disease. J Pediatr Gastroenterol Nutr 2011 Sep;53(3):280-6.

(14) Varni JW, Bendo CB, Denham J, Shulman RJ, Self MM, Neigut DA, et al. PedsQL Gastrointestinal Symptoms Module: Feasibility, Reliability, and Validity. J Pediatr Gastroenterol Nutr 2014 May 5.

(15) Lane M, Varni JW, Czyzewski DI, Shulman RJ. Health-related quality of life in pediatric functional vs. Organic gastrointestinal (GI) disorders: The pedsql gastrointestinal symptoms module. Gastroenterology 2012142(5 SUPPL. 1):S382.

(16) Wailoo A, Tosh J, Hemingway P. Use of Tumour Necrosis Factor Alpha (Tnf A) Inhibitors (Adalimumab and Infliximab) For Chrohns Disease. Decision Support Unit; 2009.

(17) National., National Institute for Health and Care Excellence. Infliximab for subacute manifestations of ulcerative colitis. 2008. Report No.: TA140.

(18) Xu X, Brandenburg NA, McDermott AM, Bazil CW. Sleep disturbances reported by refractory partial-onset epilepsy patients receiving polytherapy. Epilepsia 2006 Jul;47(7):1176-83.

(19) Selai C, Kaiser S, Trimble M, Price M. Evaluation of the relationship between epilepsy severity and utility. Value Health 2002;5(6):512-3.

(20) Selai CE, Elstner K, Trimble MR. Quality of life pre and post epilepsy surgery. Epilepsy Res 2000 Jan;38(1):67-74.

(21) Selai CE, Trimble MR, Price MJ, Remak E. Evaluation of health status in epilepsy using the EQ-5D questionnaire: a prospective, observational, 6-month study of adjunctive therapy with anti-epileptic drugs. Curr Med Res Opin 2005 May;21(5):733-9.

(22) Remak E, Hutton J, Selai C, Trimble MR, Price M J. A cost-utility analysis of adjunctive treatment with newer antiepileptic drugs in the UK. Journal of Drug Assessment 2004;7:109-20.

(23) Langfitt JT, Vickrey BG, McDermott MP, Messing S, Berg AT, Spencer SS, et al. Validity and responsiveness of generic preference-based HRQOL instruments in chronic epilepsy. Qual Life Res 2006 Jun;15(5):899-914.

(24) Trueman P, Duthie T. Use of the Hospital Anxiety and Depression Scale (HADS) in a large, general population study of epilepsy. Quality of Life Newsletter 1998;19:9-10.

(25) Craig D, Rice S, Paton F, Fox D, Woolacott N. Retigabine for the adjunctive treatment of adults with partial onset seizures in epilepsy with and without secondary generalisation: A Single Technology Appraisal TA 232. 2011.

(26) National Clinical Guideline Centre. The diagnosis and management of the epilepsies in adults and children in primary and secondary care CG137. 2012.

(27) Janssen MF, Lubetkin EI, Sekhobo JP, Pickard AS. The use of the EQ-5D preference-based health status measure in adults with Type 2 diabetes mellitus. Diabet Med 2011 Apr;28(4):395-413.

(28) Speight J, Reaney MD, Barnard KD. Not all roads lead to Rome-a review of quality of life measurement in adults with diabetes. Diabet Med 2009 Apr;26(4):315-27.

(29) Meadows K. Which DHP? Introducing DHP-1 and DHP-18. Diabetes Profile Patient Research 2014Available from: URL: http://www.diabetesprofile.com/57

(30) Tosh J, Brazier J, Evans P, Longworth L. A review of generic preference-based measures of health-related quality of life in visual disorders. Value Health 2012 Jan;15(1):118-27.

(31) Longworth L, Yang Y, Young T, Mulhern B, Hernandez AM, Mukuria C, et al. Use of generic and condition-specific measures of health-related quality of life in NICE decision-making: A systematic review, statistical modelling and survey. Health Technol Assess 201418(9):1-224. Available from: URL: http://www.journalslibrary.nihr.ac.uk/__data/assets/pdf_file/0008/108368/FullReport-hta18090.pdf

(32) National Institute for Health and Care Excellence. Type 2 Diabetes: The management of type 2 diabetes (CG87). National Institute for Health and Care Excellence; 2009. Report No.: CG87.

(33) National Institute for Health and Care Excellence Canagliflozin for the Treatment of Type 2 Diabetes Mellitus (ID554). 2013. Report No.: ID554.

(34) Bristol Myers Squibb, Astra Zeneca. Dapagliflozin for the treatment of type 2 diabetes (TA288). 2012. Report No.: TA288.

(35) Cummins E, Scott N, Rothnie K, Waugh N, Fraser C, Philip S, et al. Dapagliflozin for the treatment of type 2 diabetes TA288. Aberdeen HTA Group; 2012. Report No.: TA288.

(36) Cummins E, Royle P, Syhangdan D, Waugh N. Evidence Review: liraglutide for the treatment of type 2 diabetes TA203. 2009. Report No.: TA203.

(37) National Institute for Health and Care Excellence. Patient access scheme submission template TA301. 2009. Report No.: TA301.

(38) National Institute for Health and Care Excellence. Patient Access Scheme Submission Template TA274. 2009. Report No.: TA274.

(39) Hadi M, Gibbons E, Fitzpatrick R. A structured review of patient-reported outcome measures (PROMs) for colorectal cancer. Report to the department of health, 2010. Patient-reported Outcome Measurement Group 2010.

(40) Anderson H, Palmer MK. Measuring quality of life: impact of chemotherapy for advanced colorectal cancer. Experience from two recent large phase III trials. Br J Cancer 1998;77 Suppl 2:9-14.

(41) Siena S, Peeters M, Van CE, Humblet Y, Conte P, Bajetta E, et al. Association of progression-free survival with patient-reported outcomes and survival: results from a randomised phase 3 trial of panitumumab. Br J Cancer 2007 Dec 3;97(11):1469-74.

(42) Wilson TR, Alexander DJ, Kind P. Measurement of health-related quality of life in the early follow-up of colon and rectal cancer. Dis Colon Rectum 2006 Nov;49(11):1692-702.

(43) Doornebosch PG, Gosselink MP, Neijenhuis PA, Schouten WR, Tollenaar RA, de Graaf EJ. Impact of transanal endoscopic microsurgery on functional outcome and quality of life. Int J Colorectal Dis 2008 Jul;23(7):709-13.

(44) Doornebosch PG, Tollenaar RA, Gosselink MP, Stassen LP, Dijkhuis CM, Schouten WR, et al. Quality of life after transanal endoscopic microsurgery and total mesorectal excision in early rectal cancer. Colorectal Dis 2007 Jul;9(6):553-8.

(45) Gosselink MP, Busschbach JJ, Dijkhuis CM, Stassen LP, Hop WC, Schouten WR. Quality of life after total mesorectal excision for rectal cancer. Colorectal Dis 2006 Jan;8(1):15-22.

(46) Janson M, Lindholm E, Anderberg B, Haglind E. Randomized trial of health-related quality of life after open and laparoscopic surgery for colon cancer. Surg Endosc 2007 May;21(5):747-53.

(47) Hamashima C. Long-term quality of life of postoperative rectal cancer patients. J Gastroenterol Hepatol 2002 May;17(5):571-6.

(48) Sharma A, Sharp DM, Walker LG, Monson JR. Predictors of early postoperative quality of life after elective resection for colorectal cancer. Ann Surg Oncol 2007 Dec;14(12):3435-42.

(49) Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst 1993 Mar 3;85(5):365-76.

(50) Rowen D, Brazier J, Young T, Gaugris S, Craig BM, King MT, et al. Deriving a preference-based measure for cancer using the EORTC QLQ-C30. Value Health 2011 Jul;14(5):721-31.

(51) Hoyle M, Crathorne L, Peters J, Jones-Hughes T, Cooper C, Napier M, et al. The effectiveness and cost-effectiveness of cetuximab (mono- or combination chemotherapy), bevacizumab (combination with (non-oxaliplatin chemotherapy) and panitumumab (monotherapy) for the treatment of metastatic colorectal cancer after first-line chemotherapy (review of technology appraisal 150 and part review of technology appraisal 118): a systematic review and economic model TA242. 2011.

(52) National Clinical Guideline Centre, National. Colorectal Cancer: the diagnosis and management of colorectal cancer CG131. 2011.

(53) Whyte S, Pandor A, Stevenson M, Rees A. Bevacizumab in combination with Fluoropyrimidine-based chemotherapy for the first line treatment of metastatic colorectal cancer TA212. 2009.

(54) Merck Sorono. Erbitux (Cetuximab) for the first line treatment of metastatic colorectal cancer TA176. 2008.

(55) Sanofi UK. Aflibercept (Zaltrap®) for the treatment of metastatic colorectal cancer (mCRC) previously treated with an oxaliplatin containing regimen ID514. 2013

.

(56) Murray A, Lourenco T, de Verteuil R, Hernandez R, Fraser C, McKinley A, et al. Systematic review of the clinical effectivenesss and cost-effectiveness of laparascopic surgery for colorectal cancer TA105. 2005.

(57) Griffin S, Walker S, Sculpher M, White S, ErhornS, Brent S, et al. Cetuximab plus radiotherapy for the treatment of locally advanced squamous cell carcinoma of the head and neck TA145. 2008.

(58) Liverpool Reviews and Implementation Group. Cetuximab for recurrent and/or metastatic squamous cell carcinoma of the head and neck TA172. 2008.

(59) Peasgood T, razier J, apaioannou D. A systematic review of the validity and responsiveness of EQ-5D and SF-6D for depression and anxiety . HEDS Discussion Paper 2012;No.12.15(http://eprints.whiterose.ac.uk/74659).

(60) Gunther OH, Roick C, Angermeyer MC, Konig HH. The responsiveness of EQ-5D utility scores in patients with depression: A comparison with instruments measuring quality of life, psychopathology and social functioning. J Affect Disord 2008 Jan;105(1-3):81-91.

(61) Saarni SI, Suvisaari J, Sintonen H, Pirkola S, Koskinen S, Aromaa A, et al. Impact of psychiatric disorders on health-related quality of life: general population survey. Br J Psychiatry 2007 Apr;190:326-32.

(62) Mann R, Gilbody S, Richards D. Putting the 'Q' in depression QALYs: a comparison of utility measurement using EQ-5D and SF-6D health related quality of life measures. Soc Psychiatry Psychiatr Epidemiol 2009 Jul;44(7):569-78.

(63) Petrou S, Morrell J, Spiby H. Assessing the empirical validity of alternative multi-attribute utility measures in the maternity context. Health Qual Life Outcomes 2009;7:40.

(64) Swan J, Sorrell E, MacVicar B, Durham R, Matthews K. "Coping with depression": an open study of the efficacy of a group psychoeducational intervention in chronic, treatment-refractory depression. J Affect Disord 2004 Oct 1;82(1):125-9.

(65) Peveler R, Kendrick T, Buxton M, Longworth L, Baldwin D, Moore M, et al. A randomised controlled trial to compare the cost-effectiveness of tricyclic antidepressants, selective serotonin reuptake inhibitors and lofepramine. Health Technol Assess 2005 May;9(16):1-134, iii.

(66) Serfaty MA, Haworth D, Blanchard M, Buszewicz M, Murad S, King M. Clinical effectiveness of individual cognitive behavioral therapy for depressed older people in primary care: a randomized controlled trial. Arch Gen Psychiatry 2009 Dec;66(12):1332-40.

(67) Caruso R, Rossi A, Barraco A, Quail D, Grassi L. The Factors Influencing Depression Endpoints Research (FINDER) study: final results of Italian patients with depression. Ann Gen Psychiatry 2010;9:33.

(68) Fernandez JL, Montgomery S, Francois C. Evaluation of the cost effectiveness of escitalopram versus venlafaxine XR in major depressive disorder. Pharmacoeconomics 2005;23(2):155-67.

(69) Reed C, Monz BU, Perahia DG, Gandhi P, Bauer M, Dantchev N, et al. Quality of life outcomes among patients with depression after 6 months of starting treatment: results from FINDER. J Affect Disord 2009 Mar;113(3):296-302.

(70) Sapin C, Fantino B, Nowicki ML, Kind P. Usefulness of EQ-5D in assessing health status in primary care patients with major depressive disorder. Health Qual Life Outcomes 2004 May 5;2:20.

(71) Konig HH, Born A, Heider D, Matschinger H, Heinrich S, Riedel-Heller SG, et al. Cost-effectiveness of a primary care model for anxiety disorders. Br J Psychiatry 2009 Oct;195(4):308-17.

(72) Konig HH, Born A, Gunther O, Matschinger H, Heinrich S, Riedel-Heller SG, et al. Validity and responsiveness of the EQ-5D in assessing and valuing health status in patients with anxiety disorders. Health Qual Life Outcomes 2010;8:47.

(73) Bosmans JE, Hermens ML, de Bruijne MC, van Hout HP, Terluin B, Bouter LM, et al. Cost-effectiveness of usual general practitioner care with or without antidepressant medication for patients with minor or mild-major depression. J Affect Disord 2008 Nov;111(1):106-12.

(74) Lamers LM, Bouwmans CA, van SA, Donker MC, Hakkaart L. Comparison of EQ-5D and SF-6D utilities in mental health patients. Health Econ 2006 Nov;15(11):1229-36.

(75) van SA, Cuijpers P, Smits N. Effectiveness of a web-based self-help intervention for symptoms of depression, anxiety, and stress: randomized controlled trial. J Med Internet Res 2008;10(1):e7.

(76) Supina AL, Johnson JA, Patten SB, Williams JV, Maxwell CJ. The usefulness of the EQ-5D in differentiating among persons with major depressive episode and anxiety. Qual Life Res 2007 Jun;16(5):749-54.

(77) Sobocki P, Ekman M, Agren H, Krakau I, Runeson B, Martensson B, et al. Health-related quality of life measured with EQ-5D in patients treated for depression in primary care. Value Health 2007 Mar;10(2):153-60.

(78) Aydemir R, Ergun H, Soygur H, Kesebir S, Tulunay C. Quality of life in major depressive disorder: A cross sectional study. Turkish Journal of Psychiatry 2009;20:205-12.

(79) Ergun H, Aydemir O, Kesebir S, Soygur H, Tulunay FC. SF-36 and EQ-5D quality of life instruments in major depressive disorder patients: Comparisons of two different treatment options. Value Health 2007;10(6):A303.

(80) Mychaskiw M, Hoffman D, Dodge W. EQ-5D index scores by remission status in patients with generalized anxiety disorder. Int J Neuropsychopharmacol 2008;11:279.

(81) Fitzpatrick R, Gibbons E, Mackintosh A. An overview of patient-reported outcome measures for people with anxiety and depression. Report to the department of health, 2009. Patient-reported Outcome Measurement Group 2009.

(82) EEPRU Programme. EEPRU. www eepru org uk 2014Available from: URL: www.eepru.org.uk

(83) National Institute for Health and Care Excellence. Depression in adults. The treatment and management of depression in adults CG90. 2009.

(84) McCrone P, Chisholm D, Knapp M, Hughes R, Comi G, Dalakas MC, et al. Cost-utility analysis of intravenous immunoglobulin and prednisolone for chronic inflammatory demyelinating polyradiculoneuropathy. Eur J Neurol 2003 Nov;10(6):687-94.

(85) Kaltenthaler E, Shackley P, Stevens K, Beverley C, Parry G, Chilcott J. Computerised cognitive behaviour therapy for depression and anxiety TA51. 2002.

(86) Kuyken W, Byford S, Taylor RS, Watkins E, Holden E, White K, et al. Mindfulness-based cognitive therapy to prevent relapse in recurrent depression. J Consult Clin Psychol 2008 Dec;76(6):966-78.

(87) Scott J, Palmer S, Paykel E, Teasdale J, Hayhurst H. Use of cognitive therapy for relapse prevention in chronic depression. Cost-effectiveness study. Br J Psychiatry 2003 Mar;182:221-7.

(88) Cahill J, Stiles WB, Barkham M, Hardy GE, Stone G, Agnew-Davies R, et al. Two short forms of the Agnew Relationship Measure: the ARM-5 and ARM-12. Psychother Res 2012;22(3):241-55.

(89) Papaioannou D, Brazier J, Parry G. How valid and responsive are generic health status measures, such as EQ-5D and SF-36, in schizophrenia? A systematic review. Value Health 2011 Sep;14(6):907-20.

(90) Pitkanen A, Valimaki M, Endicott J, Katajisto J, Luukkaala T, Koivunen M, et al. Assessing quality of life in patients with schizophrenia in an acute psychiatric setting: reliability, validity and feasibility of the EQ-5D and the Q-LES-Q. Nord J Psychiatry 2012 Feb;66(1):19-25.

(91) McCrone P, Patel A, Knapp M, Schene A, Koeter M, Amaddeo F, et al. A comparison of SF-6D and EQ-5D utility scores in a study of patients with schizophrenia. J Ment Health Policy Econ 2009 Mar;12(1):27-31.

(92) Barton GR, Hodgekins J, Mugford M, Jones PB, Croudace T, Fowler D. Measuring the benefits of treatment for psychosis: validity and responsiveness of the EQ-5D. Br J Psychiatry 2009 Aug;195(2):170-7.

(93) Royal College of Psychiatrists. Outcomes Measures Recommended for use in Adult Psychiatry. 2011. Report No.: OP78.

(94) Royal College of Psychiatrists Faculty of the Psychiatry of Old Age. Individual Patient Outcome Measures Recommended for Use in Older People's Mental Health. 2012. Report No.: OP86.

(95) European Medicines Agency, Committee for Medicinal Products for Human Use. Guideline on clinical investigation of medicinal products, including depot preparations in the treatment of schizophrenia. 2012.

(96) Department of Health. The NHS Outcomes Framework 2014/15. Department of Health; 2013.

(97) National Institute for Health and Care Excellence. Aripiprazole (Abilify®) for the treatment of schizophrenia in adolescents (15-17 years) TA213. 2010.

(98) Karlawish JH, Zbrozek A, Kinosian B, Gregory A, Ferguson A, Glick HA. Preference-based quality of life in patients with Alzheimer's disease. Alzheimers Dement 2008 May;4(3):193-202.

(99) Coucill W, Bryan S, Bentham P, Buckley A, Laight A. EQ-5D in patients with dementia: an investigation of inter-rater agreement. Med Care 2001 Aug;39(8):760-71.

(100) Naglie G, Tomlinson G, Tansey C, Irvine J, Ritvo P, Black SE, et al. Utility-based Quality of Life measures in Alzheimer's disease. Qual Life Res 2006 May;15(4):631-43.

(101) Ankri J, Beaufils B, Novella JL, Morrone I, Guillemin F, Jolly D, et al. Use of the EQ-5D among patients suffering from dementia. J Clin Epidemiol 2003 Nov;56(11):1055-63.

(102) Jonsson L, Andreasen N, Kilander L, Soininen H, Waldemar G, Nygaard H, et al. Patient- and proxy-reported utility in Alzheimer disease using the EuroQoL. Alzheimer Dis Assoc Disord 2006 Jan;20(1):49-55.

(103) Thorgrimsen L, Selwood A, Spector A, Royan L, de Madariaga LM, Woods RT, et al. Whose quality of life is it anyway? The validity and reliability of the Quality of Life-Alzheimer's Disease (QoL-AD) scale. Alzheimer Dis Assoc Disord 2003 Oct;17(4):201-8.

(104) Bostrom F, Jonsson L, Minthon L, Londos E. Patients with dementia with lewy bodies have more impaired quality of life than patients with Alzheimer disease. Alzheimer Dis Assoc Disord 2007 Apr;21(2):150-4.

(105) Vogel A, Mortensen EL, Hasselbalch SG, Andersen BB, Waldemar G. Patient versus informant reported quality of life in the earliest phases of Alzheimer's disease. Int J Geriatr Psychiatry 2006 Dec;21(12):1132-8.

(106) Bryan S, Hardyman W, Bentham P, Buckley A, Laight A. Proxy completion of EQ-5D in patients with dementia. Qual Life Res 2005 Feb;14(1):107-18.

(107) European Medicines Agency. Guideline on clinical investigation of medicinal products in the treatment of Parkinson's disease . 2012.

(108) European Medicines Agency. Guideline on medicinal products for the treatment of alzheimers disease and other dementias. 2008.

(109) Wolfs CA, Dirksen CD, Kessels A, Willems DC, Verhey FR, Severens JL. Performance of the EQ-5D and the EQ-5D+C in elderly patients with cognitive impairments. Health Qual Life Outcomes 2007;5:33.

(110) Mulhern B, Rowen D, Brazier J, Smith S, Romeo R, Tait R, et al. Development of DEMQOL-U and DEMQOL-PROXY-U: generation of preference-based indices from DEMQOL and DEMQOL-PROXY for use in economic evaluation. Health Technol Assess 2013 Feb;17(5):v-140.

(111) Peninsula Technology Assessment Group. 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 TA217. 2010.

(112) Dyer MT, Goldsmith KA, Sharples LS, Buxton MJ. A review of health utilities using the EQ-5D in studies of cardiovascular disease. Health Qual Life Outcomes 2010;8:13.

(113) Schweikert B, Hahmann H, Leidl R. Validation of the EuroQol questionnaire in cardiac rehabilitation. Heart 2006;92(1):62-7.

(114) van Stel HF, Buskens E. Comparison of the SF-6D and the EQ-5D in patients with coronary heart disease. Health Qual Life Outcomes 2006;4(1):20.

(115) Nowels D, McGloin J, Westfall JM, Holcomb S. Validation of the EQ-5D quality of life instrument in patients after myocardial infarction. Qual Life Res 2005;14(1):95-105.

(116) Eurich DT, Johnson JA, Reid KJ, Spertus JA. Assessing responsiveness of generic and specific health related quality of life measures in heart failure. Health Qual Life Outcomes 2006;4(1):89.

(117) Spertus J, Peterson E, Conard MW, Heidenreich PA, Krumholz HM, Jones P, et al. Monitoring clinical changes in patients with heart failure: a comparison of methods. Am Heart J 2005;150(4):707-15.

(118) De SD, et al. "Validity and reliability of three commonly used quality of life measures in a large European population of coronary heart disease patients". International Journal of Cardiology 2013;167(5):2294-9.

(119) Sut H K, Unsar S. "Is EQ-5D a valid quality of life instrument in patients with acute coronary syndrome". Journal Anadolu Kardiyol Derg 2011;1:156-62.

(120) Garster NC, Palta M, Sweitzer NK, Kaplan RM, Fryback DG. Measuring health-related quality of life in population-based studies of coronary heart disease: comparing six generic indexes and a disease-specific proxy score. Qual Life Res 2009 Nov;18(9):1239-47.

(121) Kontodimopoulos N. "The impact of disease severity on EQ-5D and SF-6D utility discrepancies in chronic heart failure". The European Journal of Health Economics 2011;12(4):383-91.

(122) Feeny D, et al. Agreement about identifying patients who change over time cautionary results in cataract and heart failure patients. Medical Decision Making 2012;32(2):273-86.

(123) Withers KL, White J, Carolan-Rees G, Patrick H, O'Callaghan P, Murray S, et al. Patient reported outcome measures for cardiac ablation procedures: a multicentre pilot to develop a new questionnaire. Europace 2014 Mar 13.

(124) De SD, Clays E, Annemans L, Doyle F, Kotseva K, Pajak A, et al. Health related quality of life in coronary patients and its association with their cardiovascular risk profile: results from the EUROASPIRE III survey. Int J Cardiol 2013 Sep 30;168(2):898-903.

(125) De SD, Clays E, Annemans L, De BD. EQ-5D versus SF-12 in coronary patients: Are they interchangeable? Value Health 201417(1):84-89.

(126) Bristol Myers Squibb, Pfizer. Apixaban (Eliquis) for the prevention of stroke and systemic embolism in people with non-valvular atrial fibrillation TA275. 2012.

(127) BMJ Assessment Group. Rivaroxaban for the prevention of stroke and systemic embolism in people with atrial fibrillation TA256. 2011.

(128) Spackman E, Burch J, Faria R, Borbacho B, Fox D, Woolacott N. Dabigatran etexilate for the prevention of stroke and systemic embolism in atrial fibrillation . Evidence Review Group Report TA249. 2011.

(129) McKenna C, Maund E, Sarowar M, Fox D, Stevenson M, Pepper C, et al. Dronedarone for atrial fibrillation and atrial flutter TA197. 2009.

(130) Southampton Health Technology Assessment Centre. Implantable cardioverter defribillators for the treatment of arrythmias and cardiac resynchronisation therapy for the treatment of heart failure (review of TA95 and TA120) ID481. 2013.

(131) Fox M, Mealing S, Anderson R, Dean J, Stein K, Price A, et al. The effectiveness and cost-effectiveness of cardiac resynchronisation (Biventricular Pacing) for Heart Failure: A systematic review and economic model TA120. 2006.

(132) Castelnuevo E, Stein K, Pitt M, Garside R, Payne L. The effectiveness and cost-effectiveness of dual chamber pacemakers for bradycardia due to atrioventricular block or sick sinus syndrome: systematic review and economic model TA88. 2004.

(133) BMJ Technology Assessment Group. Ivabradine for the treatment of chronic heart failure TA267. 2012.

(134) National Institute for Health and Care Excellence. Review of NICE Technology Appraisal Guidance No.182; Prasugrel for the treatment of acute coronary syndromes with percutaneous coronary intervention ID648. 2010.

(135) Greenhalgh J, Bagust A, Boland A, Dwan K, Beale S, Fleeman N, et al. Prasugrel with percutaneous coronary intervention for treating acute coronary syndromes TA182. 2013.

(136) Bagust A, Boland A, Blundell M, Fisher M, Dundar Y, Greenhalgh J, et al. Ticagrelor for the treatment of acute coronary syndromes TA236. 2011.

(137) Simpson E, Fitzgerald P, Evans P, Tappenden P, Reckless J, Bakhai A, et al. Bivalirudin for the treatment of ST-segment elevation myacardial infarction TA230. 2011.

(138) Hill R, Boland A, Dickinson R, Dundar Y, Haycox A, McLeod C, et al. Drug-eluting stents: a systematic review and economic evaluation TA152. 2005.

(139) NICOR. NACSA Annual Report. 2012.

(140) Ward S, Lloyd-Jones M, Pandor A, Holmes M, Ara R, Ryan A, et al. Statins for the prevention of coronary events. 2005.

(141) Granger CB, Goldberg RJ, Dabbous O, Pieper KS, Eagle KA, Cannon CP, et al. Predictors of hospital mortality in the global registry of acute coronary events. Arch Intern Med 2003 Oct 27;163(19):2345-53.