White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/106561/

Version: Published Version

Monograph:
Ara, R., Duarte, A., Harnan, S. et al. (2015) 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 2013/14 National Clinical Audit Programme. Appendix E - Diabetes. Report. Policy Research Unit in Economic Evaluation of Health and Care Interventions (EEPRU) (RR0032). Policy Research Unit in Economic Evaluation of Health and Care Interventions

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

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?

APPENDIX E, DIABETES

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

RR0032

July 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.

INDEX

1 BACKGROUND

2 OVERVIEW

3 METHOD

- 3.1 Psychometric properties (WP1.1)
- 3.2 Alternative measures (WP1.2)
- 3.3 Evidence required for economic evaluations (WP1.3)

4 RESULTS FOR DIABETES

- 4.1 Evidence of appropriateness of EQ-5D in diabetes
- 4.2 Alternative measures in diabetes
- 4.3 Evidence for economic evaluations in diabetes
- 4.4 Recommendations for diabetes

5 SUMMARY

6 REFERENCES

7 APPENDIX

TABLES

Table 1 Selection of most appropriate review for diabetes

Table 2 Summary of evidence on EQ-5D for diabetes

Table 3 Summary of existing TAs in diabetes

Table 4 Mandatory fields collected in the core diabetes NCA

Table 5 Recommendations and associated future research for diabetes

Table 6 Summary of evidence currently available for recommended measure(s)

FIGURES

Figure 1 Modelling approach used in the diabetes HTAs

Figure 2 Modelling utilities as a continuous measure in diabetes macular oedema

Acronyms

Acronym	Definition
ADA	American Diabetes Association
ADDQoL	audit of diabetes-dependent quality of life
ADS	Appraisal of Diabetes Scale
AEs	Adverse events
AMSTAR	Assessing the quality of systematic reviews
ASA	American Society of Anesthesiologists classification,
BCVA	Best corrected visual activity
BMI	Body mass index (kg/m^2)
BPI-DPN	Brief Pain Inventory- Diabetic Peripheral Neuropathy
CES-D	Center for Epidemiologic Studies Depression
CG	Clinical guideline
CHU-9D	Child Health Utility 9D
DE	Data extraction
DH	Department of Health
DHP	Diabetes health profile
DHP-1	Diabetes health profile-1
DHP-18	Diabetes health profile-18
DR	Diabetic retinopathy
DSC-R	Diabetes Symptom Checklist – Revised
DSIS	The Daily Sleep Interference Scale
EASD	European Association for the Study of Diabetes
ECG	Electrocardiogram
EEPRU	Policy Research Unit in Economic Evaluation of Health and Care Interventions
EQ-5D	EuroQol 5 dimensions
EQ-5D-Y	EuroQol 5 dimensions youth version
ERG	Evidence review group
ESRD	End stage renal disease
FR	Future research
HbA_{1c}	Glycated haemoglobin
HDL-c	High-density lipoprotein concentrations
HRQoL	Health related quality of life
HS	Health states
HTA	Health technology assessment
HUI	Health Utility Index
HUI2	Health Utility Index mark 2
HUI3	Health Utility Index mark 3
IV	Intravenous
LDL-c	Low-density lipoprotein concentrations
mBPI-sf	Modified Brief Pain Inventory-Short Form.
MCS-12	mental component summary of the SF-12
MDT	Multi disciplinary team
MI	Myocardial infarction

MODY	Maturity onset diabetes of the young
NCA	National Clinical Audit
NEI-VFQ-25	National Eye Institute Visual Functioning Questionnaire
NHS	National Health Service
NTSS-6	The Neuropathy Total Symptom Score
PCS-12	physical component summary of the SF-12
PDPN	Painful diabetic peripheral neuropathy
PedsQL™	Paediatric quality of life inventory™
PR	Potential recommendations
PREM(s)	Patient reported experience measure(s)
PROM(s)	Patient reported outcome measure(s)
PVD	Peripheral vascular disease
QA	Quality assessment
QALYs	Quality adjusted life years
QOL-DN	quality of life in diabetes neuropathy instrument
R&D	Research and development
RCT	Randomised controlled trial
SBP	Systolic blood pressure
SD	Standard deviation
SF-6D	Short form 6D
SF-36	Short form 36
SF-12	Short form 12
SG	Standard gamble
SS	Study selection
STA(s)	Single technology appraisal(s)
T2DM	Type 2 diabetes mellitus
TA(s)	Technology Appraisal(s)
TC	Total cholesterol
TIA	Transient ischaemic attack
TTO	Time trade off
UK	United Kingdom
UKPDS	United Kingdom Prospective Diabetes Study
VAS	Visual analogue scale
VFQ-25	Visual Functioning Questionnaire 25
WP	Work package

1. BACKGROUND

The Policy Research Unit in Economic Evaluation of Health and Care Interventions (EEPRU) was approached by Jason Cox (Research and 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 Department of Health (DH) (8th November 2013). The current document provides details on the objectives, methodology and results for Work Package 1 (WP1): to determine what PROMs should be used in the 13 health conditions specified in the 2014/15 NCA programme.

2. OVERVIEW

WP1 is split into three separate components consisting of:

- WP1.1 To examine whether the EuroQol-5D (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.

This Appendix provides the results for diabetes and should be read in conjunction with both the main report and the method/search strategy appendices.

3. METHOD

The full detailed methodology used is provided in Appendix A and B, including the search strategy, selection criteria for studies included, and data extraction etc. In summary, a review of the literature was undertaken to assess the appropriateness of the EQ-5D in terms of classic psychometric criteria (WP1.1); where the EQ-5D was not considered appropriate, additional searches were undertaken to identify alternative measures (WP1.2); and finally, existing health technology assessments were reviewed and data requirements were compared with variables currently collected in the diabetes audit (WP1.3).

3.1 Psychometric properties (WP1.1)

Assessments reported in the included studies were categorised according to the following definitions:

Acceptability

Data relating to how acceptable the measure was to the person completing it, expressed as the proportion of completed surveys, or the proportion of missing data.

Reliability

There are two main definitions for reliability, a) the degree to which a measure reproduces the same results in an unchanged population and b) the degree to which a measure reproduces the same results when completed by different assessors (e.g. patient and proxy report). In both cases, reliability can be assessed by re-testing, and calculating the correlations or difference between tests. In case a) the comparison may be between the same populations separated by time, where no change in health state was observed (as compared to using an alternative condition specific or generic measure). In case b) the measure may be completed by multiple people (proxies) on the patient's behalf and their responses compared with those of the patient. Where the outcome measure is specifically designed for self-report by patients, this test of reliability may be expected to produce less agreement.

Construct validity

This is an assessment of how well an instrument measures what it intends to measure. Two main definitions are used in this review.

a) Known group validity, where estimates for groups that are known to differ in a concept of interest are compared either qualitatively or statistically. The known groups may be defined using other measures, according to clinical categorisation.

b) Convergent validity assesses the extent to which a measure correlates with other measures of the same or similar concepts. Correlation coefficients were considered low if <0.3, moderate if between 0.3 and 0.5, and strong when >0.5.

Responsiveness

a) Change over time. This is an assessment of whether measurements using the instrument can detect a change over time, where a change is expected. This may be before and after an intervention, or through progression of a disease. Evidence was considered to be good where a t-test was significant, though weaker evidence to support responsiveness was considered where there was a change in the expected direction, but was not statistically significant or not tested. Effect size and standardised response mean were also acceptable assessments of responsiveness.

b) Ceiling and floor effects were also considered to be indicators of responsiveness. Assessments of ceiling effects include the proportion of patients who score full health within a group of patients with known health detriments. A ceiling or floor effect can affect the sensitivity of the measure in detecting changes over time in patients at the extremes of the measure (for example those with severe disease activity and those with just minor symptoms of the condition).

3.2 Alternative measures (WP1.2)

Where the EQ-5D was considered appropriate, no further searches were performed.

3.3 Evidence required for economic evaluations (WP1.3)

The existing Health Technology Assessments (HTAs) were reviewed alongside the variables currently collected in the NCA to determine if clinical or PROM data routinely collected in the NCAs would suffice to address questions of cost-effectiveness, and to identify any gaps in the evidence that would be required to compare providers, or the cost-effectiveness of interventions or policies.

4. RESULTS FOR DIABETES

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

4.1.1 *Selection of systematic review*

One systematic review was identified through expert sources,(1) and two (2;3) from the Longworth et al. review.(4) The process of selection of the most appropriate review is documented in Table 1. Janssen et al. was selected as it provides more detail about the psychometric properties of the EQ-5D, and is also marginally more recent than the Oxford review.(1)

Table 1: Selection of most appropriate review for diabetes

Review	Search date	Relevance of review	Quality of search	Quality of review	Selection
Oxford (2009)(1)	September 2008	Question relevant, but too little psychometric data provided	Reliance on pre-existing database, pubmed strategy not provided. However, probably adequate.	No QA; no search numbers; single reviewer DE and SS; synthesis involved two reviewers	Exclude – less recent than Janssen, less DE detail than Janssen
Janssen et al 2011(2)	January 2009	Question relevant, some detail provided	Searched pubmed and EMBASE. Good supplemental searches.	No QA; details of search numbers provided; unclear reviewers SS, unclear DE; synthesis unclear	Include – more recent than Oxford review, more detail provided
Speight et al 2009(3)	Not Assessed	Not a psychometric study	Not Assessed	Not Assessed	Exclude – question not relevant

QA, quality assessment; DE, data extraction; SS, study selection.

4.1.2 *Structured abstract for Janssen et al 2011(2)*

Purpose of review

Amongst other objectives that are not relevant to WP1.1, Janssen et al. (2011) aimed to "review the scientific evidence on the measurement performance of the EQ-5D in the assessment of health realted quality of life (HRQoL) in adults with Type 2 diabetes (T2DM), with a focus on the ability of the EQ-5D to distinguish between different complications and levels of severity."

Methods of review

Search and study selection: EMBASE and MEDLINE (database host platform used was unclear), the EuroQoL website, and the research databases of the International Diabetes Federation (IDF), American Diabetes Association (ADA) and European Association for the Study of Diabetes (EASD) were searched. Reference lists and author's personal collections were searched. Electronic searches were conducted from 1987 to January 2009, using terms for 'EQ-5D' combined with 'diabet*' or specific diabetic complications in EMBASE and MEDLINE. However, the full search strategies were not reported and terms used in the research database searches were not given. Studies were selected for inclusion if they reported EQ-5D measurement properties or scores on the EQ-5D index, visual analogue scale (VAS) or percentage dimension scores, or the relative impacts on utilities (e.g. beta coefficients) of specific complications.

Data extraction and synthesis: Data was extracted (unclear whether double data extraction or data checking performed) using a previously developed and tested standardised form. A narrative synthesis was performed according to the psychometric quality assessed, namely validity, reliability and responsiveness. These qualities were defined as a) convergent validity, the degree to which theoretically related measures agree; b) discriminant validity, the degree to which theoretically unrelated measures do not agree (this property is not considered as a separate form of validity in WP1, but may be included in convergent validity where reported); c) known group validity, a form of construct validity where groups known to vary in health level are shown to vary by the measure; d) discriminative validity, described as a "prerequisite" to construct validity, where the measure is shown to distinguish between levels of health within or between patients (considered as equivalent to known group or convergent validity, as defined in WP1); and e) responsiveness, the ability of the measure to detect clinically meaningful changes over time.

Results of the review

Janssen et al. (2011) included 39 articles which presented evidence on the measurement properties of the EQ-5D in T2DM. The narrative synthesis was brief, and the tabulation of study results was not detailed. Convergent and discriminant validity (a & d) were reported in 14 studies, which mostly examined the strength of correlation between the EQ-5D and other generic health status measures or disease-specific measures. Construct validity (c) was reported in 16 studies. These mainly used regression and ANOVA techniques. Most demographic and clinical categories were discriminated between by the EQ-5D, with the exception of patients with multiple conditions and in patients with mild disease. A ceiling effect was noted in several studies in comparison to the short form -6D (SF-

6D). Additional properties not listed in the methods section of the review were described in the narrative, namely predictive validity (not relevant to this review), responsiveness, and reliability. The EQ-5D was shown to be responsive in studies that assessed this, with the exception of one study investigating diabetic peripheral neuropathic pain. Reliability was reported in two studies as being "good" or "excellent" using the k statistic and intraclass correlation coefficients (ICC).

Review authors' conclusions

The review authors concluded that evidence supported construct, convergent and discriminant validity, test–retest reliability and responsiveness of the EQ-5D in T2DM. However, they also noted that a ceiling effect and an inability to capture health detriments due to multiple complications were observed in several studies.

4.1.3 Assessment of review in relation to objectives of work package 1.1

Relevance of review question: One of the aims of Janssen 2011(2) is convergent with the aims of WP1.1.

Assessment of review quality: Janssen et al. (2011) scored poorly against the relevant AMSTAR criteria (Table A1, Appendix). The authors did not provide a reference to a published protocol to evidence an a priori design, meaning the study is potentially open to bias in terms of changes to the analysis plan in response to the results found. No quality assessment of the included studies was conducted, and it was therefore not possible to formulate conclusions which take the quality of the included studies into account. It does not appear that double data extraction or data-checking took place, leaving the study at higher risk of errors. There appears to have been more than one reviewer involved in study selection, but it is not clear if this constituted double-checking of study selection or just division of labour. As such, there is a small risk that some studies may have been missed in study selection. The meta-analysis conducted is not of relevance to the aims of WP1.1. In addition, theselection criteria are poorly defined.

Acceptability of the search: In addition to searching relevant electronic bibliographic databases, the review authors also searched several professional organization websites and performed reference list checking. Even though full strategies were not given, the search approach is considered adequate for the review.

Acceptability of study selection: Study selection criteria were not well defined, and reference to the full text of included articles (to retrieve additional data) revealed that studies had been included which would not have met our selection criteria.

Adequacy of available data and synthesis: The review only provided a small amount of data relating to each study, and this was not adequate for the requirements of WP1.1. Not enough detail was provided for some of the studies to enable a judgement to be made about whether the evidence supported the conclusions. The synthesis was very brief.

In conclusion, the methods employed in the review required some remedial action. Whilst the searches were thought to be adequate, the inclusion criteria appeared to be wider than that of WP1.1. In addition, the data extraction and synthesis were not detailed enough to allow a thorough understanding of the psychometric properties of the EQ-5D in this population. As such, all studies were re-considered for inclusion, and a detailed data extraction and synthesis of these studies performed.

4.1.4 Reanalysis of Janssen et al. 2011(2)

Of the 39 studies initially included in the review, 16 met the inclusion criteria of WP1.1.(6-22) Study characteristics and results are provided in Tables A2 to A8, Appendix.

In brief, one study used the USA EQ-5D tariff,(10) 13 studies used the UK EQ-5D tariff,(6-9;11;13;15-17;19-22) of which four (8;9;11;20) were also conducted in the UK, the remainder being conducted in Europe in four cases,(6;7;15;16) Australia in one case, (17) Thailand in one case,(22) Singapore in one case(142) and multinational in two cases.(6;13) The tariff used was not clear in Gore et al. 2005 (set in USA),(12) and no tariff was used in Vernon et al. 2008 (unclear setting).(14)

Patient characteristics differed somewhat across studies. Broad inclusion criteria were used in most studies, with the exception of Glasziou et al.(17) who recruited normotensive patients, four studies which recruited patients with painful neuropathy,(12-14;16) and one study which recruited patients with diabetic retinopathy.(9) Mean ages were similar across studies, ranging from 52(21) to 69(19) years old. The number of withdrawals was very poorly reported, with most studies only reporting responders. Three studies were post hoc analyses of randomised controlled trials (RCTs),(6;13;14) one study was a time series,(19) whilst the remainder were cross sectional studies. Only one study cited the measurement of psychometric properties of EQ-5D as their main aim.(10)

Acceptability: It was difficult to assess the acceptability of the EQ-5D as the majority of samples used in the studies appear to be from respondents who completed the full set of variables tested. However, one study did report that of the participants who completed the questionnaire (which comprised several measures including the EQ-5D), none of the EQ-5D items were left unanswered, indicating that the EQ-5D is acceptable.(8)

Reliability: Two studies assessed reliability (Table A7, Appendix).(20;22) Clarke et al. reported K statistics ranging from 0.59 (95% confidence interval (CI) =0.45–0.74) for the EQ-5D mobility dimension to 0.26 (95% CI =0.11–0.40) for the EQ-5D pain dimension, with a good ICC of 0.59 (95% CI 0.41 to 0.72) for the tariff scores.(20) Sakthong et al. also reported a good correlation (r=0.74, 95% CI 0.57 to 0.84, p<0.001).(22) Both studies suggest reliability is good (Appendix).(22)

Construct validity (known group): Nine studies reported known group validity for the EQ-5D in people with diabetes (Table A5, Appendix).(9;11-13;15-17;22;23) Matza et al. compared the mean EQ-5D for various dichotomised known groups. These groups included the median split for two disease-specific tools (the appraisal of diabetes scale (ADS); the diabetes symptom checklist – revised (DSC-r)), whether patients wanted to lose weight, or wanted to stay the same weight; whether patients had daytime hypoglycaemia; whether they had night-time hypoglycaemia; whether they had any hyperglycaemia; whether they were treated with oral medication or injected insulin. The mean EQ-5D was significantly different between groups in all cases by t-test, except for presence of hyperglycaemia or type of treatment. Matza et al. concluded that whilst the EQ-5D is valid, it should not be used as the sole measure in a clinical trial.(23)

In Vexiau et al. 2008, all EQ-5D health dimensions scores showed differences between those with and without hypoglycaemia symptoms, although the differences were only statistically significant for pain/discomfort and anxiety/depression (p<0.005). There was also a statistically significant difference between mean EQ-5D index scores between these two groups (0.70 (SD 0.26) vs. 0.80 (SD 0.23) respectively, p<0.0005). Glasziou et al. 2007 presented graphs that showed the mean EQ-5D deficit at baseline was significant for those with (compared to those without) stroke/transient ischaemic attack (TIA); peripheral revascularisation/amputation; myocardial infarction (MI); hospital admission for unstable angina; currently treated hypertension, but the deficit was not significant for those with diabetic eye disease, or coronary artery bypass graft. Similar graphs were shown for the SF-6D (short form 12 (SF-12)) and SF-6D (short form-36 (SF-36)), with a similar pattern of sensitivity

across the classes. In addition, these latter two measures did not produce a significant difference in means for those with or without MI, where the EQ-5D did. Sakthong et al.(22) compared means in EQ-5D index in those with and without various characteristics, and found statistically significant differences for neuropathy, retinopathy, nephropathy and cardiovascular disorder, but not for glycated hypoglycaemia (HbA_{1c}).

Known group validity of the EQ-5D for visual functioning was further assessed in one study against a condition specific measure (Visual Functioning Questionnaire 25 (VFQ-25)), and against a clinical measure (visual acuity).(9) Formal statistics were not presented, but trends in EQ-5D scores were the same at the upper and lower extremes of visual acuity, though middle range values were not well differentiated by the EQ-5D, possibly due to small sample sizes as the same problem was observed with the Health Utility Index-3 (HUI-3) and VFQ-25.

Four studies compared neuropathic pain scales or generic pain scales to the EQ-5D. These were the neuropathy total symptom score-6 questionnaire (NTSS-6),(11) the quality of life in diabetes neuropathy instrument (QOL-DN),(11) the brief pain inventory modified for pain in diabetic peripheral neuropathy (BPI-DPN),(12) and the modified brief pain inventory short form (mBPI-sf)(13;16). Three studies recruited only patients with neuropathic pain,(12;13;16) whilst one recruited anyone with diabetes.(11) All found good agreement, which supported the EQ-5D's ability to detect pain related to diabetic neuropathy, even in a general sample of patients with diabetes.

Construct validity (convergent): Seven studies considered convergent validity of the EQ-5D compared to a variety of other measures (Table A6, Appendix).(6-8;14;17;21;22) Convergent validity between the EQ-5D index and the WHO Diabetes treatment satisfaction questionnaire (WHO-DTSQ) was low in the study that assessed treatment satisfaction ($r=0.28$ $p<0.0001$).(7) The EQ-5D does not have an item related to treatment satisfaction, so this is perhaps not surprising.

Two studies compared EQ-5D against a general diabetes scale. One study used two such scales: the appraisal of diabetes scale (ADS) and the diabetes symptom checklist- revised (DSC-R).(24) The study found moderate to strong agreement (r range: -0.44 to -0.61 (all $p<0.001$)), except for the ophthalmic scale of the DSC-R, which had only low agreement ($r=-0.22$, $p<0.05$). The second study used the audit of diabetes-dependent quality of life (ADDQoL),(21) and showed a strong correlation (spearman's rank=0.54) between the EQ-5D index score and the scores of those who scored well on

the ADDQoL for the items relating to current QoL (rather than items relating to overall impact of diabetes on life domains).

The EQ-5D was compared against the daily sleep interference score (DSIS) in one study(14) which recruited only those with painful peripheral neuropathy. This study reported low to moderate agreement (r range: 0.08 to 0.44) for the EQ-5D, with moderate correlation only being observed in the pain/discomfort dimension, suggesting that the impact of painful neuropathy on sleep was not captured in the EQ-5D.

The correlation of the EQ-5D with HbA_{1c} was assessed by two studies.(6;22) Sakthong et al. reported a low but significant correlation (spearman's rank r=-0.17, p<0.01) between the EQ-5D and HbA_{1c} at baseline.(22) In Bech et al. changes in the HbA_{1c} did not have a significant correlation with changes in the EQ-5D in response to treatment.(6) Both studies recruited people with T2DM with similar HbA_{1c} scores at baseline (7.7% (SD 1.7, range 4.0–15.8) vs. 7.7% (SD 1.7, range not reported) respectively).

Other diabetic complications were also used in convergent validity tests. The EQ-5D was correlated with the SF-6D in a ranking of severity of seven complications (r=0.837 to SF-6D (SF12), and r=0.842 to SF-6D (SF36)).(17) It was also shown to have moderate correlations with the number of complications (r=-0.40, p<0.01), and a measure of depression, the Center for Epidemiologic Studies - Depression (CES-D) (-0.49, p<0.01), but low correlation with body mass index (BMI) (r=-0.15, p<0.01)(22)

Overall, convergent validity was generally strong when compared with generic or condition-specific measures. There were, however, certain situations where the correlations were low. These include some uncertainty (owing to small sample numbers) about the ability of the EQ-5D to detect ophthalmic issues,(8), low or non-significant correlations with HbA_{1c} in the two studies that reported this comparison, low correlations with a measure of treatment satisfaction(7) and low correlations with BMI in two studies.(8;22) On balance, results suggest a lack of responsiveness of the EQ-5D to changes in HbA_{1c}, rather than a complete lack of correlation.(6) There were also lower correlations with the daily sleep interference scale (DSIS)(14) and with a measure of treatment satisfaction in diabetes (WHO-DTSQ).(7)

Responsiveness (change over time): Two studies assessed responsiveness through changes over time (Table A8, Appendix). Bech et al.(6) reported that the EQ-5D showed little or no change, as

expected, over a four month period, although the WHO-DTSQ did. Johnson et al.(19) recorded mean EQ-5D values each year over three years. They reported the expected decreases in EQ-5D mean scores over time (within subject effects analysis for time and diabetes status: f=4.49 p=0.012), supporting the responsiveness of the EQ-5D.

Responsiveness (ceiling effects): Three studies assessed responsiveness through examining potential ceiling effects on the EQ-5D (Table A8, Appendix).(8;10;21) Matza et al. 2007 noted that 40% of those with diabetes scored full health on the EQ-5D, whilst 0% scored full health on the Psychological General Well-Being Index. Similarly, Bharmal & Thomas 2006 noted that amongst 165 people with diabetes (from a general population sample) who scored full health on the EQ-5D, the mean physical component summary of the SF-12 (PCS-12) was significantly different to those with no medical conditions, whereas the mental component summary of the SF-12 (MCS-12), and total score of the SF-6D were not significantly different. This indicates a ceiling effect in the EQ5D in comparison to the PCS-12. In Sakthong et al. 37.8% of people with diabetes reported full health, and of these, the mean ADDQoL was -3.4 (SD 2.49), indicating a ceiling effect of the EQ-5D in comparison with the ADDQoL.(22)

4.1.5 Conclusion of appropriateness of EQ-5D in diabetes

The evidence base used to assess the appropriateness of the EQ-5D in patients with diabetes is relatively large (N=16), with the majority using data obtained using the UK EQ-5D tariff, and all using adults 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 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 2).

Table 2: 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.				

N: number of studies; KGV: known group validity

4.2 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.(5) 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.(25) 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.(4;26) 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.(4) 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. However, the authors caution that sample sizes were small, and that further research with larger sample sizes is required.

There will be some children in the diabetes audit but it is thought that these will be in the minority due to the age-related prevalence of the condition. Consequently, due to the time constraints of the project, the evidence base describing potential alternative PROMs for paediatrics with diabetes was not reviewed.

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

4.3.1 Cost-effectiveness modelling approach used in recent HTAs in diabetes

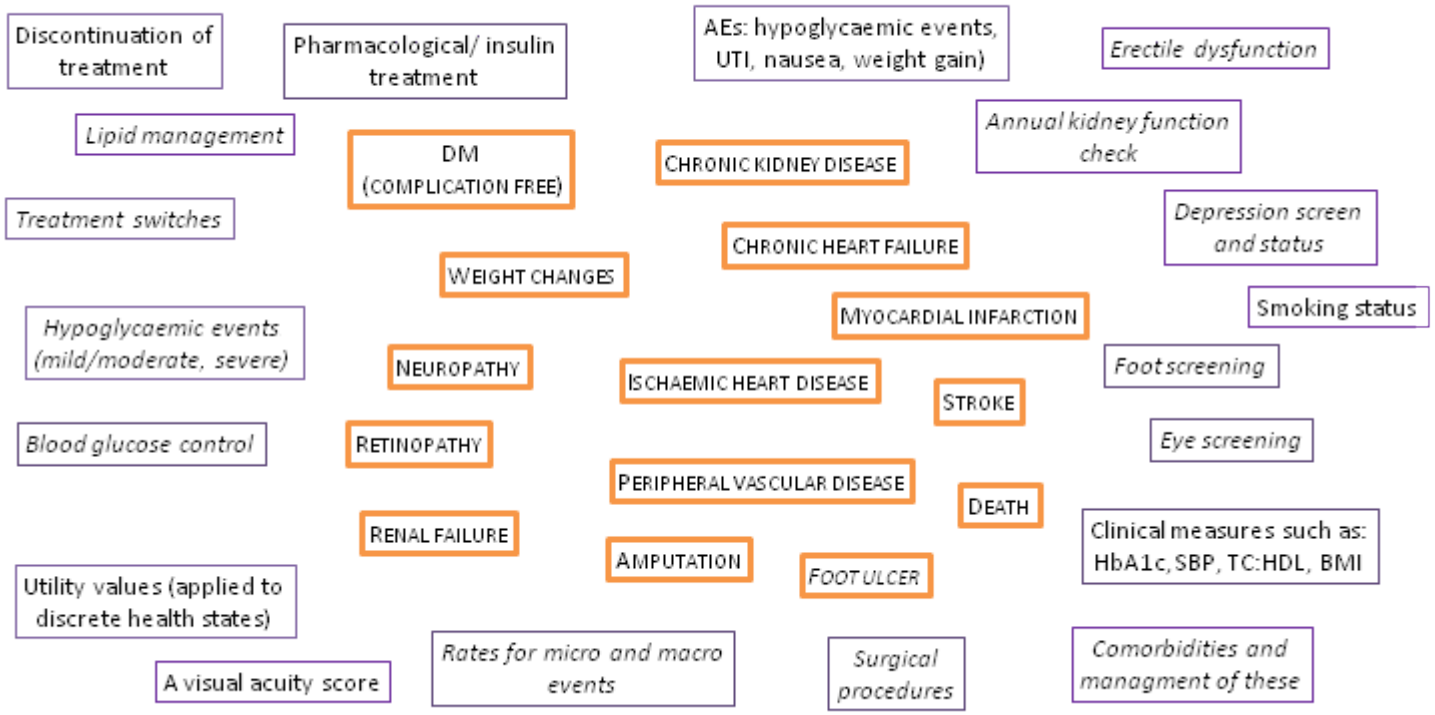
Ten technology appraisals (TAs) relating to diabetes were identified from the searches. A CG was subsequently identified from the references lists of the included studies.(27) Two of the TAs were superseded by rapid reviews of the evidence,(28;29) one was withdrawn (see clinical guideline (CG) 87),(30) and one was suspended due to licence withdrawal.(31) All the TAs were in T2DM. The CG and four single technology appraisals (STAs) compared insulin therapies,(27;32-35) and the remaining two STAs compared interventions for diabetic related macular oedema.(36;37) One examined the clinical and cost-effectiveness of a pharmaceutical intervention,(28;29) while the other examined the clinical and cost-effectiveness of an intravitreal implant (Table 3).(28;29)

The models comparing insulin therapies were generally constructed around the United Kingdom Prospective Diabetes Study (UKPDS), using individual patent level simulation Markov models comprising of discrete health states representing micro and macrovascular diabetic complications (Figure 1).(20) Clinical trial data were used together with UKPDS risk functions to describe the clinical effects of the interventions. UKPDS risk functions are available for congestive heart failure, ischaemic heart disease, myocardial infarction, stroke, blindness, ulcer, amputation, and renal failure.(20) The variables required to use the functions included: HbA_{1c}, BMI, systolic blood pressure (SBP), high-density lipoprotein (HDL-c), low-density lipoprotein (LDL-c), white blood cell, haemoglobin, heart rate, epidermal growth factor receptor, presence of micro/macro albuminuria, atrial fibrillation and peripheral vascular disease. Markov models were used for both macular oedema interventions and the discrete health states were defined using severity grades based on best corrected visual activity (BCVA),(36) or visual acuity scores (Figure 2).(37) Effectiveness of the interventions was modelled using clinical trial data which provided evidence of changes in either BCVA or visual acuity scores.

All studies quality adjusted survival by assigning mean utility values to the discrete health states. The models comparing insulin interventions used EQ-5D evidence predominantly sourced from the UKPDS.(20) Exceptions were the disutilities associated with weight changes and hypoglycaemic events (which are not included in the UKPDS). Conversely, the utilities used in the models comparing interventions for diabetic macular oedema were modelled using regressions which mapped from the clinical variables (BCVA, visual acuity scores) to HRQoL data (EQ-5D, non-societal preferences) as shown in Figure 2.(38) The blue (diamond) line and red (square) line show the

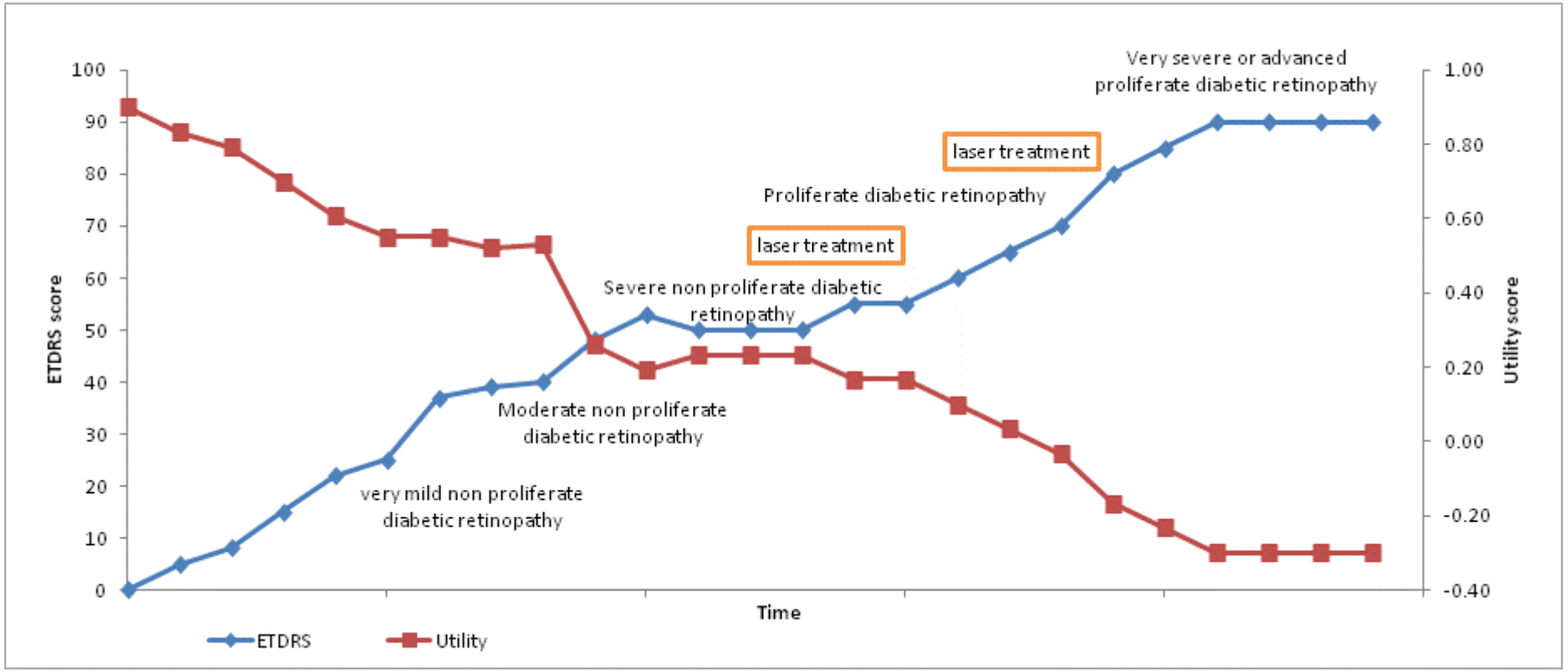
changes in disease severity (measured using visual acuity scores) and utilities (measured using EQ-5D) over time respectively.

Figure 1: Modelling approach used in the diabetes HTAs



Legend: Orange framed boxes with uppercase text describe the health states used in the diabetes TA models while the purple framed boxes with lower case (non italic) text describe the evidence used. The boxes with italised text are additional information which would ideally be collected to inform future economic models.

Figure 2: Modelling utilities as a continuous measure in diabetes macular oedema



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. The EQ-5D may not be sensitive to changes in sight and thus not appropriate for this condition (as discussed above).(4;26)

Table 3: Summary of existing TAs in diabetes

Model approach	Method used to model utilities
STA: Diabetes (type 2) - canagliflozin; 2014(32)	
Patient level Markov model Health states: complication free, chronic kidney disease (7 stages), neuropathy (5 conditions), retinopathy (4 conditions), ischemic heart disease, myocardial infarction, chronic heart failure, stroke, peripheral vascular disease, death. Additional variations: weight changes (BMI), hypoglycaemic events (mild, moderate, severe), upper and lower UTIs, GMI, gastrointestinal upset (nausea) Effectiveness: principally driven by HbA_{1c}, BMI, SBP, cholesterol (used in UKPDS functions); rates for AEs (hypoglycaemic events, UTI, GMI, nausea) and discontinuation; retinopathy Source: clinical RCTs (network meta-analyses)	Utility: EQ-5D supplemented with non-preference data (TTO); mean values assigned to discrete HS Source: published literature (CODE-2 study plus others) AEs: discrete HS utilities cover majority of these, additional changes in utility modelled for changes in BMI (statistical relationship)
STA (TA288): Type 2 diabetes - dapagliflozin combination therapy; 2013(34)	
Patient level DES, predominantly UKPDS Health states: ischemic heart disease, myocardial infarction, chronic heart failure, stroke, amputation, blindness, end stage renal disease, death. Additional: weight changes, hypoglycaemic events Effectiveness: HbA_{1c}, SBP, TC:HDL, BMI (used in UKPDS functions); rates for AEs (hypoglycaemic events, UTI, GMI, nausea) and discontinuation Source: clinical RCTs	Utility: EQ-5D, HUI, supplemented with expert opinion; mean values assigned to discrete HS for macro & micro HS Source: published literature AEs: discrete HS utilities cover majority of these, additional changes in utility modelled for changes
STA(TA248: Diabetes (type 2) - exenatide (prolonged release); 2012(39)	
Markov model Health states: predominantly UKPDs Effectiveness: HbA_{1c}, SBP, TC:HDL (UKPDS functions) Effectiveness: Source: clinical RCTs	Utility: EQ-5D, HUI, assumption; mean values assigned to discrete HS Source: published literature AEs: HRQoL loss due to treatment induced nausea, reduced fear of hypoglycaemic episodes
CG (CG87): Type 2 diabetes: newer agents for blood glucose control (update of CG66), 2010(27)	
Patient level simulation (UKPDS outcomes model) Health states: ischemic heart disease, myocardial infarction, heart failure, stroke, amputation, blindness, renal failure, hypoglycaemic events Effectiveness: HbA_{1c}, SBP, TC:HDL (UKPDS functions) Source: clinical RCTs	Utility: EQ-5D supplemented with assumptions for weight change; mean values assigned to discrete HS Source: UKPDS (Clarke) and literature AEs: HRQoL change due to treatment induced weight change, QoL loss due to treatment induced nausea
STA (TA203): Diabetes (type 2) - liraglutide; 2010(35)	

Markov model Health states: micro & macrovascular complications, hypoglycaemic events Effectiveness: HbA_{1c}, SBP, TC:HDL, BMI Source: clinical RCTs	Utility: EQ-5D predominantly; mean values assigned to discrete HS Source: Clarke et al UKPDS, supplemented by published literature AEs: HRQoL change due to treatment induced weight change, HRQoL loss due to hypoglycaemic events
STA (TA301)[a]: Diabetic macular oedema - fluocinolone acetonide intravitreal implant; 2013(36)	
Markov model used to extrapolate beyond the RCT duration (i.e. at 3 years) 13 discrete health states defined by 5 ETDRS bands of the BCVA in the treated eye, plus death Effectiveness: improvement in ETDRS criteria Source: clinical RCTs	Utility: TTO and SG from patients with AMD; mean values assigned to discrete HS Source: TTO exercise with 72 patients with AMD Values are for the BCVA in the best seeing eye AEs: disutility due to AEs (cataract development, raised intraocular pressure) not modelled
STA (TA274)[b]: Macular oedema (diabetic) - ranibizumab; 2013(36)	
Markov model (cohort) 8 discrete health states defined by visual acuity scores (0-25; 26-35; 36-45; 46-5; 56-65; 66-75; 76-85; 86-100 letters) plus death Effectiveness: change in visual acuity Source: clinical RCTs	Utility: EQ-5D; regression between visual acuity scores and EQ-5D; mean values assigned to discrete HS Source: published literature AEs: disutility due to treatment toxicity not modelled

HS: health states; AEs: Adverse Events; STA: Single Technology Appraisal; TA: Technology Appraisal; CG: Clinical Guideline; TTO: Time trade-off; SG: Standard Gamble, RCT: randomised controlled trial; BCVA: best corrected visual activity; AMD: age-macular degeneration; PLS: patient level simulation; GMI: genital mycotic infection; UTI: urinary tract infection; ETDRS: Early Treatment of Diabetic Retinopathy Study eye chart, DES: discrete event simulation; TC: total cholesterol; BMI: body mass index; SBP: systolic blood pressure; UKPDS: United Kingdom Prospective Diabetes Study; HbA_{1c}: glycated haemoglobin; HUI; health utility index; CG: clinical guidelines; TC:HDL: ratio of total cholesterol to high density lipoprotein

[a] details on modelling approach taken from TA271 as TA301 review this model and do not construct a new model. [b]details on modelling approach taken from TA237 as TA274 review this model and do not construct a new model

Diabetes is a complex disease and patients are at risk of multiple diabetes related complications. The existing economic evaluations described above cover just a proportion of the possible health states in a typical clinical pathway. Areas not explored or modelled in detail include (for example) leg ulcers, erectile dysfunction, surgical interventions, annual screening for complications, day to day management of diabetes, management of diabetes when hospitalised for comorbidities etc. Many of these may have implications in terms of comparing providers or policies. Although not intended to be exhaustive, examples of the evidence required to extend the economic approach beyond what is explored in the existing models are provided in Figure 1 above.

The following core evidence would be required to compare providers or the cost-effectiveness of interventions for diabetes:

- Blood glucose control
- Clinical variables used in the UKPDS functions (HbA_{1c}, SBP, ratio of total cholesterol to high density lipoprotein (TC:HDL), BMI)

- Both micro and macro vascular events (e.g. ischaemic heart disease, chronic heart failure, renal failure, peripheral vascular disease)
- Surgical procedures (type of intervention (e.g. revascularisation, amputation), success rate, post-surgical complication, length of stay etc)
- Pharmaceutical interventions (type of intervention, concomitant medications, adverse events)
- Screening uptake and results (foot, eye, etc)
- Utility values (collected alongside condition severity and surgical interventions)
- Death rates (diabetes related, all cause)

The majority of this evidence would need to be linked through timings of collection.

4.3.2 Fields collected in the core diabetes NCAs

The National Diabetes Audit integrates data from both primary and secondary care sources from all patients (irrespective of age) diagnosed with diabetes (all types of diabetes mellitus, excluding gestational diabetes, impaired glucose tolerance, impaired glucose fasting) in England and Wales. Participation in the audit is voluntary for primary care but all trusts with specialist diabetes services in England and Wales are expected to participate in the audit. It is understood that the audit is currently expanding to gather information on: pregnancy care in women with diabetes, and diabetes footcare, and these are discussed below.(40)

Details of the fields in the core diabetes NCA are provided in two documents. There are very few mandatory fields (NHS number, NHS organisation code, diabetes type, date of diagnosis, GP practice code, sex, date of birth, postcode), and records which do not collect the full complement of mandatory fields are rejected (Table 4).(40) However, in the optional fields, there is an exhaustive list of codes for different diabetes diagnoses with and without diabetes related complications, and a list of fields providing clinical parameter levels such as HbA_{1c}, cholesterol, BMI etc. (Table A9, Appendix). If these can be obtained from GP patient records automatically, this increases the evidence available from the audit considerably.

There is also a National Diabetes Inpatient Audit, which includes a Patient Experience questionnaire and an associated Bedside Audit form (Table A9, Appendix). The information collected provides a snapshot (from a pre-specified day) of diabetes inpatients' perceptions of their experiences relating

to the support and services received. The proposed objective of this audit is to use the results to improve inpatient experience. The bedside audit collects information (completed by the nurse) relating to the patient's diabetes type and treatment received. The patient experience questionnaire (completed by the patient) collects information relating to staff seen, ability to provide their own diabetes care (insulin and testing of blood sugar levels), the appropriateness and timeliness of the food provided and the patient's perception of the diabetes related support from hospital staff.

Table 4: Mandatory fields collected in the core diabetes NCA

PATIENT DEMOGRAPHIC/OBSERVATION DATA[a]
NHS number, Date Of birth, Postcode (patient's usual address), Gender
PROVIDER INFORMATION[a]
GP practice code, NHS organisation code (provider code)
CLINICAL HISTORY[a]
Date of diagnosis (diabetes), Diabetes type (e.g. Type 1, Type 2, MODY, other specified, not specified)
OBSERVATIONS[a]
No mandatory fields (Multiple clinical variables in optional data e.g. BMI, BP, Cholesterol, HbA_{1c} etc. See Appendix)
2 CODES [b]
Diabetes mellitus diagnosis (plus multiple combinations of complications, e.g. DM with no complication, DM with hyperosmolar coma, DM with renal manifestation, DM with ophthalmic manifestation, DM with peripheral circulatory disorder; Insulin dependent DM with gangrene, Insulin dependent DM with nephropathy, Insulin dependent DM with hypoglycaemic coma, Insulin dependent DM with diabetic cataract, Non-insulin Insulin dependent DM with ulcer, T2DM with multiple complications etc) Latest diagnosis code of diabetes mellitus (multiple combinations as above)

NHS, nation health service: GP: general practitioner; MODY: maturity onset diabetes of young; DM, diabetes mellitus; BMI, body mass index; BP, blood pressure; HBA_{1C}, glycated haemoglobin; T2DM, type 2 diabetes mellitus

[a]National Diabetes audit CSV Specification 2012-2013 V5.0 01/05/2013; [b]NDA Primary care extraction specification 2012-2013 (linked to primary care records unless patient dissents)

4.3.3 Comparing fields in diabetes NCA with variables used in existing HTAs

The existing HTA models comparing insulin therapies use the UKPDS risk functions to model the benefits of treatments in terms of reductions in both micro and macrovascular complications. The key variables (HbA_{1c}, BMI, SBP) required for the functions are noted as optional fields in the current core diabetes NCA (Appendix). While many of the variables required are not currently listed, they may be available from GP records via the primary care audits. However, it is not clear if a key variable (frequency and severity of hypoglycaemic events) is recorded anywhere.

The existing HTA models comparing interventions for diabetes related visual acuity use a clinical grading measure (e.g. 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. However, it is understood that there are two additional new components currently being piloted. The Patient Experience of Diabetes Services, will measure patients' experiences in primary care and specialist services (initial results due June 2014), and the National Diabetes Foot Care Audit will explore: if nationally recommended care structures for management of diabetic foot disease are in place, if the treatment of active diabetic foot disease complies with national guidance, and if the outcomes of diabetic foot disease are optimised (due to launch Summer 2014). It is possible that these audits will enhance the existing fields with information directly relating to patient experience. However, 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.

While it is possible that many of the utility values required for economic evaluations will be available from the literature, the inclusion of a preference-based HRQoL measure (preferably the EQ-5D) in the diabetes audits would be useful for gaps in the evidence base such as HRQoL associated with hypoglyceamic events and vision. For patients with diabetes related visual acuity, it would be beneficial to include a variable (such as the VFQ-25, BCVA, or visual acuity score) to grade this condition, and potentially the EQ-5D vision bolt-on. The variable used should be selected on the basis that it could ultimately be used to weight survival to generate quality adjusted life years (QALYs) (for example via a mapping mechanism). Finally, with the exception of attendance on the DESMOND programme, the core NCA does not include any information on interventions or procedures received. However, both the current patient experience questionnaire and the diabetes inpatient audit contain numerous questions relating to diabetic complications, control, prescribing and drug management errors, intravenous insulin infusions, involvement of specialist diabetes teams, and general foot care, all of which would be useful information for economic evalutions.

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. With the exception of the two additional audits currently being piloted, the existing diabetes inpatient audit, and the patient experience of diabetes services audit, no ongoing or scheduled research in the area of PROMs for the diabetes NCA are known.(40)[personal communication, Eleanor Bunn, Audit coordinator, 13th May, 2014]

4.4 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 diabetes 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 discussed below. All suggested future research areas are indicative and would require a discussion and detailed proposal if required.

As the NCA for diabetes covers both primary and secondary care, this will involve patients across the full spectrum of the condition, from newly diagnosed patients to patients with long standing diabetes with complications such as end stage renal disease. While there is no suggestion that the EQ-5D suffers from a floor effect in this patient group, the suggested ceiling effect could be problematic for newly diagnosed patients with no diabetic related complications. It is recommended that consideration is given to the inclusion of the EQ-5D in all the different diabetes audits (i.e. adapting the current inpatient 'patient experience questionnaire' for use in the primary care setting) if possible to capture HRQoL scores across the full spectrum of the condition (PR.1). As mentioned in previous sections, the use of the EQ-5D-5L, could potentially reduce any ceiling effect in patients less severely affected by the condition. The psychometric properties of this instrument would need to be assessed in patients with diabetes (FR.1). This study would require the concurrent collection of a measure against which the EQ-5D could be compared, together with additional information such as patient demographs, diabetic related complications (micro and macrovascular), hypoglycaemic events, clinical variables (e.g. HbA_{1c}) and current medications etc. It is recommended that the DHP (see Section 3.2) is collected alongside the EQ-5D to capture issues relating to daily management of the condition (PR.11).

A potential solution to possible issues in capturing changes in HRQoL in patients with vision problems (such as diabetes related macular oedema), might be the use of a bolt-on to the EQ-5D (PR.12). The 'bolt-on' methodology involves including a question relating to vision as an addition to the standard five questions in the EQ-5D. Exploratory research eliciting preferences using time trade-off (TTO) methods from a sample of the UK population demonstrated this methodology could potentially have a significant effect upon EQ-5D valuations. However, due to limitations such as the relatively small sample size, the authors recommend additional research in larger samples is required (FR.1).(4) On the same issue, the concurrent collection of a HRQoL measure and a clinical

variable such as the VFQ-25, BCVA or visual acuity score would enhance this evidence in terms of usefulness in future economic models (PR.2). Some evidence exists which could potentially be used to link these measures to utilities, depending on which clinical measure and which preference-based measure was used in the audit.(41;42)

As discussed previously, the DH PROMS use the DHP which captures the day to day issues relating to managing diabetes. Hypoglycaemic events are a common consequence of insulin and weekly rates have been estimated at 0.82 and 0.33 for Type 1 and insulin-treated T2DM respectively.(43) Hypoglycaemic episodes can range from benign (remedied by eating fast-acting carbohydrates), to seizure, coma and even death.(44) Severe or frequent hypoglycaemic events can be traumatic for patients with diabetes. Preference-based HRQoL data from people who experience these events, and the impact on HRQoL associated with the fear of a future hypoglycaemic events are particularly sparse (Section 6.2). A study using data collected in the NCA exploring these issues would add considerably to the existing evidence base in this area and would inform future economic models in the UK and wider settings (FR.3).

Although in the minority, it is believed there will be some paediatrics in the diabetes audit. As in the previous sections, it is recommended that consideration is given to the inclusion of paediatric preference-based HRQoL questionnaires (PR.4), which again would require a primary piece of research to assess the psychometric properties in children with diabetes (FR.4).

While it is understood that the primary and secondary care audits could potentially be linked, and that additional detail from the GP records could be obtained, this is by no means clear. In addition, it is not clear if these would suffice to provide information on variables such as HbA_{1c}, BMI, SB, HDL-c, LDL-c, white blood cell, haemoglobin, etc. which are the key clinical variables used in cost-effectiveness models in diabetes. It is also not clear if it would be possible to link these to current treatment. A thorough inspection of the audit data would answer many of these queries and enable a more robust assessment of what would be required to perform economic evaluations with the current audit data (FR.5).

Table 5: Recommendations and associated future research for diabetes

PR.1	*Include the new version of the EQ-5D (EQ-5D-5L) and the DHP in future adult patient questionnaires*
FR.1	*Assess the psychometric properties of the EQ-5D-5L and the DHP in adults with diabetes using data collected in the audit*
PR.2	*Include the vision bolt-on to the EQ-5D for patients with vision problems*
FR.2	*Conduct a study to generate preference-weights for the EQ-5D vision bolt-on*
PR.3	*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.3	*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.4	*Include paediatric preference-based HRQoL instruments (e.g. Child Health Utility 9D (CHU-9D) and the HUI2 or Paediatric quality of life inventory™ (PedsQL)) in future paediatric questionnaires*
FR.4	*Assess the psychometric properties of the paediatric preference-based tools in paediatrics with diabetes using data collected in the audit*
FR.5	*Detailed analyses of fields currently collected in the diabetes audit to identify recommendations for future mandatory fields*

5. SUMMARY

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

A reanalysis of an existing review (n=16 primary studies) provided evidence that the acceptability and reliability of the EQ-5D are good (Table 6). 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.

Table 6: Summary of evidence supporting the psychometric properties of EQ-5D in all conditions

Condition	Measure	N	Acceptability	Reliability	Construct		Responsiveness		Overall
					KGV	Convergent	Change over time	Ceiling Effect	
Diabetes	EQ-5D	16	Good	Good	Good	Mixed	Good	Poor	Acceptable*
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						

N: number of studies used to inform conclusions; KGV: known group validity; *Not appropriate for DM related vision problems, or neuropathy

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

The existing audit does not include a patient questionnaire. A patient reported experience measure (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.

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 routine 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. 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.

APPENDIX: DIABETES

The tables in this Appendix provide additional information for the reviews (WP1.1, 1.2 and 1.3) conducted for diabetes.

Table A1: Quality assessment of Janssen et al. review of EQ-5D in diabetes(2)

Quality assessment criteria	**Compliance with criteria**
AMSTAR	
Was an a priori design provided?	Yes
Was there duplicate study selection and data extraction?	Unclear
Were the methods used to combine the findings of the studies appropriate?	Unclear
Was the scientific quality of the included studies assessed and documented?	Unclear
Was the scientific quality of the included studies used appropriately in formulating conclusions?	Unclear
Overall judgement of quality of review	Mostly unclear quality
Quality of the searches	Acceptable
Strength of the evidence	
Were the conclusions robust and conclusive?	No, evidence was mixed and limited
Quantity of the evidence	
Was there enough data to be confident that any additional data published subsequently would be very unlikely to change the conclusions drawn?	yes
Adequacy of data reported	
Did the review provide sufficient data to allow integration of an update/assessment of the methods used?	No
Did the review assess EQ-5D in a way compatible with the aims of work package 1.1?	No, wider inclusion criteria, lack of clarity about how psychometrics properties assessed

Table A2: Characteristics of studies included in the reanalysis of Janssen et al. for diabetes(2)

Study ref Author, Year	Country	Disease/treatment stage	Treatment (if any)	Study type (e.g. cross sectional, RCT, cohort)	Study objective
Bech et al. 2003(6)	Multi-national Australia, Croatia, the Czech Republic, France, Greece, Israel, Macedonia, Poland, Russia, Slovenia and Spain.	Pharmacotherapy naïve Type 2 Diabetes	Repaglinide for prandial glucose regulation	Placebo controlled RCT	To assess the differential impact of the prandial glucose regulating oral hypoglycaemic drug, repaglinide, and placebo upon perceptions of quality of life (QoL) and treatment satisfaction in pharmacotherapy-naive patients with Type 2 diabetes
Bharmal and Thomas 2006(10)	USA	General population, with subgroup of patients with diabetes	NR	Cross section	The purpose of this analysis was to compare the EQ-5D and the SF-6D derived from the SF-12 to examine any ceiling effects in the EQ-5D and the SF- 6D descriptive systems in the US general population.
Clarke et al. 2002(20)	UK	Type 2 diabetes	NR	Cross section	The aim of this study was to analyze quality-of-life data from the United Kingdom Prospective Diabetes Study (UKPDS) to estimate the impact of diabetes-related complications on utility-based measures of quality of life.
Currie et al. 2007(11)	UK	Type 1 or type 2 diabetes	NR	Cross section	The aim of this study was to characterize accurately DPN symptom severity in people with diabetes and correlate this with healthcare resource use, thus financial costs, in the UK.
Glasziou et al. 2007(17)	Australia	Normotensive patients with type 2 diabetes	Various treatment regimens	Cross section (using patients in	The purpose of this study is to compare summary statistics of the estimated utility values produced by different algorithms for common complications of diabetes. In

Study ref Author, Year	Country	Disease/treatment stage	Treatment (if any)	Study type (e.g. cross sectional, RCT, cohort)	Study objective
				current RCT)	particular we are interested to see if there are systematic differences in both the absolute mean utility values and the deviations associated with each type of diabetes-related complication.
Gore et al. 2005(12)	USA	Patients with physician-diagnosed diabetic distal symmetrical sensorimotor polyneuropathy with painful symptoms (burning, prickling, tingling, and/or shooting pain in toes, feet, legs, and/or hands) of at least three months' duration.	NR	Cross section	The aim was to evaluate pain severity, pain-related interference with function, sleep impairment, symptom levels of anxiety and depression, and quality of life among patients with PDPN.
Hoffman et al. 2008(13)	19 countries across 3 regions of the world: Asia (Indonesia, Malaysia, Philippines, Republic of Korea, Singapore, Taiwan, and Thailand), Latin America (Argentina, Brazil, Chile, Colombia, Ecuador, Mexico, and Venezuela) and the Middle East (Jordan, Lebanon, Saudi Arabia, Turkey, and United Arab Emirates).	Type 1 or 2 diabetes with painful symmetrical sensorimotor DPN for >12 months, <5 years, and a pain score of at least 40 mm on the 100 mm visual analog scale (VAS) of the Short Form McGill Pain Questionnaire both at screening and randomization, and a score of at least 4 on a 0 to 10 numerical rating scale of average pain in the week prior to baseline.	NR	Post hoc analysis of RCT	To understand the human impact of painful DPN on patients in Asia, Latin America, and the Middle East, we analyzed baseline data taken from patients in a clinical study prior to receiving treatment.
Johnson et al. 2000b(19)	USA	Random selection of general population	NA	Time series (Annually, over 3 years)	To compare the health related quality of life of people with diabetes to those without chronic conditions
Lloyd et al. 2008(9)	UK	Patients with diabetic retinopathy Patients with type 1 or 2 diabetes, but no retinopathy	NR	Cross section	The study was designed to elicit preferences regarding different severities of retinopathy from people with DR, people with diabetes with no retinopathy

Study ref Author, Year	Country	Disease/treatment stage	Treatment (if any)	Study type (e.g. cross sectional, RCT, cohort)	Study objective
		Members of the UK general population			but who face the prospect of DR in the future, and a group of members of the general public. In addition, patient groups completed two generic HRQL measures and a vision-specific measure [National Eye Institute Visual Functioning Questionnaire (NEI-VFQ-25)] to also document the burden of DR.
Matza et al. 2007a(24)	UK	Type 2 diabetes	Diet/exercise: 10.8% Oral meds: 65.4% IV meds ± other treatment: 23.8%	Cross section	To validate two generic measures in patients with T2DM.
Redekop et al. 2002(7)	Netherlands	Type 2 diabetes	Approx. 67% oral treatment Others used diet/exercise or insulin	Cross section	To estimate the health-related quality of life (HRQOL) and treatment satisfaction for patients with type 2 diabetes in the Netherlands and to examine which patient characteristics are associated with quality of life and treatment satisfaction.
Sakthong et al 2008(22)	Thailand	Type 2 diabetes outpatients, aged ≥ 18 years	NR	Cross section	to examine the differences and agreements between these three weights, psychometric properties including test-retest reliability, convergent and known-groups validity, and the impact of differences in the EQ-5D scores on the outcome of cost-utility analysis in Thai people.
Tolle et al. 2006(16)	France, Germany, Italy, Netherlands, Spain, United Kingdom	Patients with painful DPN	Various, including antidepressants, sedatives, analgesics,	Cross section	To determine the patient burden of painful diabetic peripheral neuropathy (DPN) with respect to pain intensity and impact on patient functioning and to

Study ref Author, Year	Country	Disease/treatment stage	Treatment (if any)	Study type (e.g. cross sectional, RCT, cohort)	Study objective
			antiepileptics.		characterize relevant DPN treatment patterns.
Vernon et al. 2008(14)	Unclear, 4 different RCTs	Diagnosis of Type 1 or 2 diabetes mellitus; diagnosis of diabetic, distal, symmetrical, sensorimotor polyneuropathy for one to five years with hemoglobin A1c levels of ≤11%; and at the baseline and randomization visits, a score of ≥40 mm on the Visual Analog Scale (VAS) of the Short-Form McGill Pain Questionnaire (SF-MPQ).	NR	Post hoc analysis of 8 RCTs	To evaluate the psychometric characteristics of the Daily Sleep Interference Scale (DSIS) in patients with painful diabetic peripheral neuropathy (DPN) or postherpetic neuralgia.
Vexiau et al. 2008(15)	France	Type 2 diabetes, who were ≥35 years old and who had been treated with metformin and a sulphonylurea for at least 6 months	Treated with metformin and a sulphonylurea	Cross section	To examine patient-reported experience of hypoglycaemia, worry about hypoglycaemic symptoms and the impact of hypoglycaemia on patients' quality of life associated with use of sulphonylurea co-administered with metformin.
Wee et al. 2006(21)	Singapore	English-speaking with type 1 or type 2 diabetes, aged ≥ 18 years	NR	Cross section	To evaluate and validate the ADDQoL questionnaire in English-speaking patients with diabetes in Singapore

PDPN, painful diabetic peripheral neuropathy; DPN, diabetic peripheral neuropathy; IV, intravenous; RCT, randomised controlled trial; UKPDS, United Kingdom Prospective Diabetes Study; PDPN, painful diabetic peripheral neuropathy; VAS, visual analogue scale; ADDQoL, audit of diabetes-dependent quality of life

Table A3: Patient characteristics of studies included in the reanalysis of Janssen et al. for diabetes[Janssen et al. 2011]

Study ref Author, Year	**Number of participants recruited**	**Age in years mean (sd); range**	**Male %**	**Ethnicity %**	**Other characteristics %**	**Missing data (patients completing study) include reasons for non-completion if given**
Bech et al. 2003(6)	253 Repaglinide: 164 Placebo: 89	Repaglinide: 56.9 (8.6); 40 to 81 Placebo: 57.3 (8.2); 40 to 76	57	Rapaglinide, Placebo African: 0.6, 0 Caucasian: 98.2, 98.9 Asian: 0, 0 Other: 1.2, 1.1	**Duration of diabetes** mean (SD) years Rapaglinide: 2.77 (4.2) Placebo: 2.81 (4.96)	Withdrawals Rapaglinide: 17.1% Placebo: 29.2%
Bharmal and Thomas 2006(10)	Diabetes: 165	Mean NR	53	White: 85 African American: 10 Asian or Pacific Islander: 4 Native American or Eskimo: 9	NR	NA (only included respondents with the full set of variables required)
Clarke et al. 2002(20)	124 in this analysis	62.3	NR	NR	NR	NR
Currie et al. 2007(11)	1,298	64 (NR)	56	NR	Complication: type 1; type 2 Acute MI: 5.5; 5.4 Stroke: 5.1; 4.2 Amputation: 2.1; 0.7 PVD no amputation: 5.1; 3.6 PVD with amputation: 1.3; 0.7 Suffer/ed leg ulcer: 13.6; 9.8 ESRD: 11.4; 6.2 Retinopathy: 22.9; 14.6 Severe loss of vision: 1.3; 0.1	NA
Glasziou et al. 2007(17)	978	67 (range 55 to 89)	71	**NR**	NR	NR
Gore et al.	265	61.3 (12.8)	45	African American/Black:	Type of diabetes	0 patients

Study ref Author, Year	Number of participants recruited	Age in years mean (sd); range	Male %	Ethnicity %	Other characteristics %	Missing data (patients completing study) include reasons for non-completion if given
2005(12)				18.8 Asian: 2.4 Caucasian: 53.3 Hispanic/Latino: 9.8 Native American: 0.4 Pacific Islander: 0 Other: 0.4 Multi-racial: 3.1 Missing: 11.8	Type I: 12.5 Type II: 86.3 Missing: 1.2	
Hoffman et al. 2008(13)	401	57 (10)	39	White: 29.7 Black: 3 Asian: 51.6 Other: 15.7	NR	NR
Johnson et al. 2000b(19)	Diabetics, n=85 No chronic condition, n=463	Diabetics: 69 (10.9) No chronic condition: 48.1 (12.7)	Diabetics: 78.4 No chronic condition: 75.1	NR	NR	NR
Lloyd et al. 2008(9)	DR, Diabetes no DR, general public: 122, 49, 150	DR; Diabetic no DR; general public: 62.2 (12.6); 52.6 (15.2); 44.4 (15.9)	50	Ethnicity: DR; Diabetic no DR; general public White: 79; 44; 72 Asian/ Asian British: 11; 25; 8 Black/Black British: 8; 23; 10 Chinese/other: 0; 0; 4	NR	NR
Matza et al. 2007a(24)	132 recruited 130 analysed	55.7 (10.3)	65	White: 80.8 Black: 6.2 Indian: 6.2 Other 6.9	Hypertension: 36.9 Diabetic retinopathy: 5.4 Depression/other mental health condition: 13.1 Other health condition: 43.1	2 unable to complete full set of questionnaires – reason not reported.

Study ref Author, Year	Number of participants recruited	Age in years mean (sd); range	Male %	Ethnicity %	Other characteristics %	Missing data (patients completing study) include reasons for non-completion if given
					None: 34.6	
Redekop et al. 2002(7)	1,371	64.9	49.8	NR	**Complications** microvascular: 22 macrovascular: 15 micro- and macrovascular: 16	Missing data for EQ-5D: 16%
Sakthong et al 2008(22)	303	61.6 (11.4)	29	NR	Neuropathy: 40.9 Retinopathy: 16.8 Nephropathy: 8.3 Cardiovascular disorders: 14.5	NR
Tolle et al. 2006(16)	140	65.6 (11.2)	58.3	NR	NR	NR
Vernon et al. 2008(14)	1,124	59	58	White: 92 Black: 4 Hispanic: 3 Asian: <1 American Indian/Alaskan: <1	Type 1: 11.6 Type 2: 88.4	NR
Vexiau et al. 2008(15)	400	62.1 (10.7)	53.6	NR	**Complications** Macrovascular:: 19.9 Microvascular: 8.3 Major medical events: 8.3	NR
Wee et al. 2006(21)	136 analysed	52	55.3	Chinese: 49 Indian: 33.8 Malay: 12.6	**NR**	Of 173 respondents, 37 excluded: 3 did not complete ADDQoL 6 completed by a caregiver 12 missed >6 items on ADDQoL 2 missed items on EQ-5D 12 skipped EQ-VAS 2 missed EQ-VAS and items on EQ-5D

NA, not applicable; NR, not reported; MI, myocardial infarction; ESRD, end-stage renal disease; SD, standard deviation; MI, myocardial infarction; PVD, peripheral vascular disease; DR, diabetic retinopathy; ADDQoL, audit of diabetes-dependent quality of life; EQ-VAS, EuroQol visual analogue scale

Table A4: Characteristics of psychometric analyses of studies included in the reanalysis of Janssen et al. for diabetes (2)

	GENERIC MEASURES			OTHER MEASURES USED			
Study ref Author, Year	**Descriptive system**	**Tariff used**	**Mean (SD); 95% CI**	**Condition-specific HRQL measures**	**Clinical measures**	**Qualitative questions**	**Other generic**
Bech et al. 2003(6)	EQ-5D	UK (Dolan)	Baseline Rapaglinide: 0.82 Placebo: 0.83	WHO-DTSQ	HbA_{1c}	None	WHO-WBQ
Bharmal and Thomas 2006(10)	EQ-5D	USA (Shaw 2005)	1 (recruited those in full health)	None	None	None	SF-6D
Clarke et al. 2002(20)	EQ-5D	UK	NR	NA	NA	NA	NA
Currie et al. 2007(11)	EQ-5D	Unclear	Type1: median 0.656 (IQR 0.248 – 0.848) Type 2: median 0.691 (IQR 0.516 – 0.796)	NTSS-6	None	None	None for analysis
Glasziou et al. 2007(17)	EQ-5D	UK	0.801 (0.206)	None	Serious adverse events	None	None
Gore et al. 2005(12)	EQ-5D	Unclear	0.5 (0.3)	BPI-DPN	None	None	None for analysis
Hoffman et al. 2008(13)	EQ-5D	UK	0.44 (0.34)	mBPI-sf	None	None	None for analysis
Johnson et al. 2000b(19)	EQ-5D	UK	Diabetics 0.72; 0.69 to 0.75 No chronic condition: 0.96 (0.94 to 0.97)	None	None	None	None
Lloyd et al. 2008(9)	EQ-5D	Unclear	NR	NEI-VFQ-25	VA		HUI-3
Matza et al. 2007a(24)	EQ-5D	UK (Dolan) Krabbe 2003	0.75 (0.3)	ADS DSC-R	None	None	PGWB
Redekop et al. 2002(7)	EQ-5D	UK (Dolan)	0.74 (0.27)	WHO-DTSQ	None	None	None
Sakthong et al. 2008(22)	EQ-5D	UK (Dolan)	0.76 (95% CI 0.74–0.78)	CES-D	HbA1c BMI Neuropathy Retinopathy	None	None

	GENERIC MEASURES			OTHER MEASURES USED			
Study ref Author, Year	**Descriptive system**	**Tariff used**	**Mean (SD); 95% CI**	**Condition-specific HRQL measures**	**Clinical measures**	**Qualitative questions**	**Other generic**
					Nephropathy Cardiovascular disorders		
Tolle et al. 2006(16)	EQ-5D	UK	NR	mBPI-SF	None	None	None
Vernon et al. 2008(14)	EQ-5D	None	NR	Sleep Interference Score	None	None	None for analysis
Vexiau et al. 2008(15)	EQ-5D	UK	0.77 (0.24) score 0.79 (0.22) weighted score		Hypoglycaemia symptoms (any, mild, moderate, severe)	none	none
Wee et al. 2006(21)	EQ-5D	UK	0.812 (range: –0.008–1) (n=148)	ADDQoL	None	none	none

WHO-DTSQ, WHO Diabetes treatment satisfaction questionnaire; WHO-WBQ, WHO wellbeing questionnaire; HbA_{1c}, glycated haemoglobin; PGWB, psychological general wellbeing index; ADS, appraisal of diabetes scale; DSC-R, diabetes symptom checklist-revised; NEI-VFQ-25, National Eye Institute Visual Functioning Questionnaire; VA, visual acuity; NTSS-6, Neuropathy Total Symptom Score-6 questionnaire; QOL-DN, Quality of Life Questionnaire–Diabetic Neuropathy; BPI-DPN, brief pain inventory modified for pain in diabetic peripheral neuropathy; mBPI-sf, modified brief pain inventory short form; NA, not applicable; ADDQoL, audit of diabetes-dependent quality of life; CES-D, Centre for epidemiologic studies – Depression.

Table A5: Know group validity data for studies included in the reanalysis of Janssen et al. for diabetes (2)

Author, Year	Method of measuring validity, Type of validity, how (e.g. known group/convergent)?	Validity results, Group A(n) vs. Group B(n)' Mean EQ-5D; mean difference in EQ-5D	Authors' conclusions/notes	Our additional conclusions/notes
Matza et al. 2007a(24)	Known group, t-test	Comparison group (n): Mean EQ-5D, t value **ADS Median Split** ADS Score ≤ 16 (n = 58): 0.88, 5.1*** ADS Score > 16 (n = 72): 0.65 **DSC-R Median Split** DSC-R Total Score ≤ 0.7(n = 61): 0.91, 6.6*** DSC-R Total Score > 0.7 (n = 69): 0.61 **Preference for Weight Change** Lose weight (n = 113):0.73, -4.4*** Stay same (n = 16): 0.92 **Experienced Hypoglycemia during the day** Yes (n = 53): 0.68, -2.6* No (n = 74): 0.82 **Experienced Hypoglycemia During the Night** Yes (n = 23): 0.60, -3.2** No (n = 101): 0.80 **Experienced Hyperglycemia** Yes (n = 64): 0.73, -1.1 No (n = 63): 0.78 **Type of Treatment** Injectable (n = 31): 0.65, -1.8 Oral only (n = 85): 0.78	The current study provides initial data suggesting that the EQ-5D and PGWB are appropriate for use in patients with type 2 diabetes, and future research may provide additional support for this conclusion.	This study provides evidence of acceptability of EQ-5D (100% completion), convergent validity (moderate to large correlations with majority of ADS and DSC-R scales ($p<0.001$)), known-group validity (comparing EQ-5D scores for sub-groups categorised by ADS score and DSC-R total score ($p<0.001$). However, there was some evidence of a ceiling effect (self-care: 92% reported no problem; preference-based index: 40% reported full health), and potential issues with the ophthalmology (relatively small correlation EQ-5D preference-based index and DSC-R ophthalmology dimensions, $r=0.22$ $p<0.05$)

Author, Year	Method of measuring validity, Type of validity, how (e.g. known group/convergent)?	Validity results, Group A(n) vs. Group B(n)′ Mean EQ-5D; mean difference in EQ-5D	Authors' conclusions/notes	Our additional conclusions/notes
Vexiau et al. 2008(15)	Known group, % scoring a problem in each group, Chi Squared Mean EQ-5D in those with and without hypoglycaemia, t-test	EQ-5D dimension: % with hypoglycaemia who scored problem; % without hypoglycaemia who scored problem, p value (Chi squared) Mobility problems: 26.7; 20.1, p=0.140 Self-care problems: 9.6; 4.6, p=0.056 Usual activities problems: 21.3; 14.0, p=0.064 Pain/discomfort problems: 59.6; 39.5, p=0.0002 Anxiety/depression problems: 58.1; 41.1, p=0.0013 EQ-5D summary score: 0.70 (0.26); 0.80 (0.23), p<0.0005		All EQ-5D health dimensions scores showed differences between those with and without hypoglycaemia symptoms although the differences were only statistically significant for Pain/discomfort and Anxiety/depression (p<0.005). EQ-5D showed significant difference between those with and without hypoglycaemia symptoms
Glasziou et al. 2007(17)	Known group, interpretation of confidence intervals on graph	Graph presented. EQ-5D mean deficit at baseline significant for those with: Stroke and/or TIA Peripheral revascularization and/or amputation MI Hospital admin for unstable angina Currently treated hypertension Not significant for: Diabetic eye disease including blindness in either eye		This study provides evidence that the EQ-5D can detect differences in utility values for diabetic complications (Spearman rank between EQ-5D and SF6D: 0.837 for vSF12; 0.842 for vSF36), and shows a differences in changes in utilities over time when comparing patients who do not have a serious adverse event and those who do.

Author, Year	Method of measuring validity, Type of validity, how (e.g. known group/convergent)?	Validity results, Group A(n) vs. Group B(n)′ Mean EQ-5D; mean difference in EQ-5D	Authors' conclusions/notes	Our additional conclusions/notes
		Coronary artery bypass graft Similar graphs shows the same pattern of significance for the SF-6D (12), whilst the SF-6D (36) fails to find a significant difference for MI in addition to diabetic eye disease and coronary artery bypass graft.		
Sakthong et al. 2008(22)	Known group, z statistic, Mann-Whitney U tests.	Comparison: mean vs. mean; difference; z statistic HBA1c<7% vs.>7%: 0.79 vs. 0.75; 0.04; -1.91 With vs. without neuropathy: 0.81 vs. 0.69; 0.12**; -5.94 With vs. without retinopathy: 0.78 vs.0.69; 0.09*; -2.16 With vs. without nephropathy: 0.77 vs. 0.67; 0.10*; -2.57 With vs. without cardiovascular disorder: 0.78 vs. 0.65; 0.13**; -3.48		Significant difference seen for neuropathy, retinopathy, nephropathy and cardiovascular disorder, but not for HBA1c.
Lloyd et al. 2008(9)	Known group, no formal comparison statistics, compared trend mean scores for decreasing level of visual acuity	Levels of visual acuity range (N): EQ-5D single index; HUI-3; VFQ-25 total score Diabetic no retinopathy (49): 0.83 ± 0.20; 0.81 ± 0.20; 90.6 ± 13.1 6/6–6/9 VA (68): 0.75 ± 0.23; 0.78 ± 0.22; 86.3 ± 13.6 6/12–6/18 VA (13): 0.50 ± 0.30; 0.30 ± 0.38; 61.5 ± 25.4 6/24–6/36 VA (10): 0.68 ± 0.29; 0.61 ± 0.35; 61.1 ± 22.6 6/60–6/120 VA (7): 0.53 ± 0.47; 0.52	No conclusion about psychometrics drawn.	The analyses show the EQ-5D is able to detect a difference in mean utility for sub-groups with different levels of impairment as defined by levels of visual acuity. The trend follows those observed in the HUI-3 and the VFQ-25 total score. Although the differences are not significant (p-value not reported), this is possibly due to the small sample sizes.

Author, Year	Method of measuring validity, Type of validity, how (e.g. known group/convergent)?	Validity results, Group A(n) vs. Group B(n)′ Mean EQ-5D; mean difference in EQ-5D	Authors' conclusions/notes	Our additional conclusions/notes
		± 0.50; 39.5 ± 24.3 Counting fingers–hand motion (3): 0.34 ± 0.36; 0.37 ± 0.00; 29.2 ± 16.1 No formal comparison statistics, but can observe that the mean EQ-5D scores decrease as the visual acuity (VA) gets worse. Very worse VA has much lower EQ than least worse VA , which is as expected. But the EQ-5D values in the middle VA subgroups do not follow same trend. This anomaly is observed in the HUI3 and the VFQ-25 data and is likely due to the very small sample sizes in the sub-groups (N: 3-49)		
Currie et al. 2007(11)	Known group, comparison of mean scores for severity by NTSS-6 scale, Kruskal-Wallis H test	NTSS-6 score category (n): EQ-5D mean score; SF-36 mean score; QOL-DN mean score 0 (335): 0.81;59.92; 25.84 >0≤3.33 (199): 0.63; 41.78; 34.76 >3.33≤7.64 (196): 0.52; 36.54; 40.83 >7.64 (202): 0.25; 25.54; 48.06 Kruskal–Wallis H-test, p<0.001 for EQ-5D and QOL-DN ANOVA, p<0.001 for SF-36	No conclusion about psychometrics drawn.	EQ-5D was able to detect significant differences in utilities across sub-groups categorised by NTSS-6-SA score, and show a similar trend in mean scores observed in the SF36 global scores, including a larger difference between groups at the extremes of the sub-groups.
Gore et al. 2005(12)	Known group, mean EQ-5D score by pain category	BPI-DPN pain category: EQ-5D mean score (sd) Mild: 0.7 (0.2) Moderate: 0.5 (0.3)		EQ-5D was able to detect significant differences in mean utilities for sub-groups categorised by BPI-DPN score (p<0.01).

Author, Year	Method of measuring validity, Type of validity, how (e.g. known group/convergent)?	Validity results, Group A(n) vs. Group B(n)′ Mean EQ-5D; mean difference in EQ-5D	Authors' conclusions/notes	Our additional conclusions/notes
		Severe: 0.2 (0.3) ANOVA f 44.7734, p< 0.00001		
Hoffman et al. 2008(13)	Known group, EQ-5D mean score by pain category, pairwise comparison	mBPI-sf pain category: EQ-5D mean score (SD) Asia ANOVA f value 28.1, p<0.0001 Mild:0.68 (0.06) Moderate: 0.61 (0.24) Severe: 0.27 (0.36) Latin America ANOVA f value 8.8, p<0.0003 Mild: 0.40 (0.31) Moderate: 0.54 (0.28) Severe: 0.28 (0.34) Middle east ANOVA f value 4.1, p<0.019 Mild: 0.67 (0.09) Moderate: 0.50 (0.31) Severe: 0.36 (0.36)		When sub-grouping by mBPI-sf Average Pain severity, EQ-5D (UK tarrif) was able to detect a statistically sig difference (p<0.05) in 3 different populations
Tolle et al. 2006(16)	Known group, EQ-5D scores according to Pain Severity Index Categories, ANOVA	EQ-5D scores according to Pain Severity Index Categories (mBPI-sf scores): Mild (1-3): 0.59 Moderate (4-6): 0.43 Severe (7-10): 0.20, P <0.001		EQ-5D mean scores decrease for subgroups categorised by Pain severity index (mild, moderate, severe) and the data seem to suggest a negative correlation between the EQ-5D and the Pain Interference data, as would be expected.

*p < 0.05; **p < 0.01; ***p < 0.001

ANOVA, analysis of variance; ADS, appraisal of diabetes scale; DSC-r, diabetes symptom checklist – revised; VFQ-25, visual functioning questionnaire; HUI-3, Health utilities index -3; NTSS-6, Neuropathy Total Symptom Score-6 questionnaire; QOL-DN, Norfolk Quality of Life Questionnaire-Diabetic Neuropathy; BPI-DPN, brief pain inventory

Author, Year	Method of measuring validity, Type of validity, how (e.g. known group/convergent)?	Validity results, Group A(n) vs. Group B(n)' Mean EQ-5D; mean difference in EQ-5D	Authors' conclusions/notes	Our additional conclusions/notes

modified for pain in diabetic peripheral neuropathy; mBPI-sf, modified brief pain inventory short form

Table A6: Convergent validity data for studies included in the reanalysis of Janssen et al. for diabetes (2)

Author, Year	Method of measuring validity, Type of validity, how (e.g. known group/convergent)?	Validity results, Group A(n) vs. Group B(n)' Mean EQ-5D; mean difference in EQ-5D	Authors' conclusions/notes	Our additional conclusions/notes
Bech et al. 2003(6)	Convergent, correlation between EQ-5D and HbA_{1c} , test unclear	There were no significant correlations between change in EQ-5D and change in HbA_{1c}.		
Redekop et al. 2002(7)	Convergent, correlation between EQ-5D and WHO-DTSQ, Pearson correlation	Correlation between WHO-DTSQ and EQ-5D Pearson r=0.28 p<0.0001		EQ-5D does not have a question directly relating to treatment satisfaction and thus EQ-5D unlikely to detect a change, unless the satisfaction has an indirect impact on HRQoL
Matza et al. 2007a(24)	Convergent validity, correlation between EQ-5D score and DSC-R subscales, Spearman correlation	ADS: -0.52*** Total Score DSC-R: -0.64*** Fatigue: -0.61*** Cognitive: -0.46*** Pain: -0.51*** Sensory: -0.53*** Cardiology: -0.56*** Ophthalmology: -0.22* Hypoglycemia: -0.44*** Hyperglycemia: -0.46*** BMI: -0.27**	Correlations between the EQ-5D index and the DSC-R subscales ranged from -0.44 to -0.61 (all p < 0.001), except for the ophthalmology subscale (r= -0.22, p<0.05)	
Vernon et al. 2008(14)	Convergent, correlation between EQ-5D and DSIS, spearman correlation coefficients.	Correlation with Daily Sleep Interference Score at baseline; 12 weeks	Small to moderate correlations	

Author, Year	Method of measuring validity, Type of validity, how (e.g. known group/convergent)?	Validity results, Group A(n) vs. Group B(n)' Mean EQ-5D; mean difference in EQ-5D	Authors' conclusions/notes	Our additional conclusions/notes
		Mobility: 0.14; 0.19 Self care: 0.13; 0.08 Usual activities: 0.23; 0.25 Pain/discomfort: 0.35; 0.44 Anxiety/depression: 0.26; 0.24		
Glasziou et al. 2007(17)	Convergent validity, EQ-5D to SF-6D, Spearman's rank	Correlations between utility measures on ranking of severity of seven complications of diabetes: EQ-5D to SF-6D (SF12): 0.837 EQ-5D to SF-6D (SF36): 0.842		There was a strong correlation between the EQ-5D and both the SF6D
Wee et al. 2006(21)	Convergent validity, EQ-5D to ADDQoL, Spearman's rank	Correlation between scores for those scoring better QoL on ADDQoL and EQ-5D score, spearman rank correlation = 0.54		The EQ-5D correlates with the ADDQoL
Sakthong et al. 2008(22)	Convergent validity, EQ-5D to other measures, Spearman's rank	Spearman's rho correlation coefficients, p value Duration of diabetes: -0.14, p<0.01 BMI: -0.15, p<0.01 HBA1c: -0.17, p<0.01 Number of diabetic complications: -0.40, p<0.01 CES-D score: -0.49, p<0.01	Based on Colton's [colton 1974] criteria, EQ-5D had small to medium correlations with duration of diabetes, BMI, HBA1c, number of complications and CES-D.	

WHO-DTSQ, WHO Diabetes treatment satisfaction questionnaire; HbA_{1c}, glycated haemoglobin; ADS, appraisal of diabetes scale; DSC-R, Diabetes symptom checklist – revised; BMI, body mass index; ADDQoL, audit of diabetes-dependent quality of life; CES-D, Centre for epidemiologic studies – Depression; DSIS, daily sleep interference score

Table A7: Reliability data for studies included in the reanalysis of Janssen et al. for diabetes (2)

Author, Year	Method of measuring validity, Type of validity, how (e.g. known group/convergent)?	Validity results, Group A(n) vs. Group B(n)′ Mean EQ-5D; mean difference in EQ-5D	Authors' conclusions/notes	Our additional conclusions/notes
Clarke et al. 2002(20)	Test-retest reliability, 4 month interval, ICC	For the 5 dimensions of the EQ-5D, the κ statistics ranged from 0.59 (95% confidence interval [CI] = 0.45–0.74) for the mobility dimension to 0.26 (95% CI = 0.11–0.40) for the pain dimension. The ICC was 0.59 (95% CI = 0.41–0.72) for the tariff scores, and therefore fell into the category of "good".	Reliability was "good".	
Sakthong et al. 2008(22)	Test-retest reliability, one week interval, ICC	Tested at one and two weeks. ICC = 0.74 (95% CI 0.57 to 0.84), p<0.001		According to Rosner 2000,[Rosner 200] agreement is good.

N, number; ICC, intraclass correlation coefficient

Table A8: Responsiveness data for studies included in the reanalysis of Janssen et al. for diabetes [Janssen et al 2011]

Author, Year	Method of measuring validity	Validity results	Authors' conclusions/notes	Our additional conclusions/notes
Bech et al. 2003(6)	Responsiveness, change over time	EQ-5D detected no change over time in response to treatment, where the WHO-DTSQ did. WHO-WBQ also detected no change.	Ceiling effect in EQ-5D "A generic utility index like the EQ-SD may be too insensitive to be used in the setting of patients with baseline wellbeing scores close to the normal population."	
Johnson et al. 2000b(19)	Responsiveness, change over time, ANCOVA	Within subject effects analysis for time and diabetes status: f 4.49 (p=0.012) Year: diabetic mean (95% CI); no chronic condition mean (95% CI) 1997: 0.72 (0.69 to 0.75); 0.96 (0.94 to 0.97) 1998: 0.68 (0.65 to 0.61); 0.96 (0.94 to 0.98) 1999: 0.65 (0.69 to 0.75);0.94 (0.93 to 0.96) When adjusted for age: f 3.14 (p=0.044)	Significantly greater decrease in HRQoL among people with diabetes compared to those without diabetes with no chronic conditions.	
Matza et al. 2007a(24)	Floor/Ceiling effect	Measure: % at floor, % at ceiling EQ-5D: 0%, 40% PGWB: 0%, 0%	Although all participants in the current sample had type 2 diabetes, 40% of the participants had the maximum EQ-5D index score of 1 which theoretically represents perfect health status. This ceiling effect suggests that the brief EQ-5D may not reflect the	

Author, Year	Method of measuring validity	Validity results	Authors' conclusions/notes	Our additional conclusions/notes
			health-related problems of all patients with type 2 diabetes, particularly patients whose symptoms have an impact on functional domains other than the five EQ-5D dimensions.	
Bharmal and Thomas 2006(10)	Ceiling effect	For the 165 patients with DM who reported no impairment on EQ-5D index (EQ-5D=1): Mean (SE) PCS-12: 52.28 (0.42), P<0.001 compared to those with no medical conditions MCS-12: 55.59 (0.42), NS SF-6D: 0.885 (0.0073), NS	Ceiling effect of EQ-5D	
Wee et al. 2006(21)	Ceiling effect	37.8% respondents did not report any problems on the EQ-5D Those who reported full health on the EQ-5D had a mean ADDQoL of -3.4 (SD 2.49)		

WHO-DTSQ, WHO Diabetes treatment satisfaction questionnaire; WHO-WBQ, WHO wellbeing questionnaire; ANCOVA, analysis of covariance; PGWB, Psychological General Well-Being Index; 95% CI, 95% confidence interval; DM, diabetes mellitus; SE, standard error; PCS-12, physical component summary of the SF-12; MCS-12, mental component summary of the SF-12; ADDQoL, audit of diabetes-dependent quality of life

Table A9: Optional fields collected in the diabetes NCA (WP1.3)

PATIENT DEMOGRAPHIC/OBSERVATION DATA[a]
Ethnic category, Death date
PROVIDER INFORMATION[a]
No optional fields
SURGERY[a]
Provider organisation, ASA grade, Cancer treatment curability, Date of surgery, Surgical urgency mode of operation, Consultant, Primary procedure, Surgical access
CLINICAL HISTORY[a]
No optional fields
OBSERVATIONS[a]
Person observation and dates for: BMI, Systolic blood pressure, Diastolic blood pressure, HbA1c level, Serum creatinine level, Urinary albumin level (and testing method), Albuminuria stage (normoalbuminuria, microalbuminuria, macroalbuminuria), Total serum cholesterol level, Diabetes routine reviews and dates for: Eye examination, Foot examination; Smoking status, Patient education review and date, Diabetes structured education programme (DESMOND) offered and date offered, Diabetes structured education programme attended and date attended.
2 CODES [b]
Persistent proteinuria, Ischemic heart disease diagnosis and date of diagnosis[*]
PATIENT EXPERIENCE QUESTIONNAIRE (Inpatient Audit 2013, completed by patient)
Section A. Background Information
Since being admitted to hospital, have you been visited by any specialist diabetes staff?
While in hospital, did a nurse or doctor make a specific effort to examine your bare feet?
Has anyone asked you about your usual diabetes medications pre-admission?
Has anyone asked you about how well your diabetes is controlled at present?
Has anyone asked about which health care professional looks after your diabetes care?
Did anyone explain that your diabetes treatment may have to change because of your admitting condition?
Were you involved in the planning of your diabetes treatment as much as you would have liked?
Have hospital staff taken your treatment preferences into account when caring for you?
Have you been able to take control of your own diabetes care while in hospital to the extent that you would like?
Do you think that the hospital staff caring for you know that you have diabetes?
Do you take insulin for your diabetes?
Are you allowed to administer insulin yourself while in hospital?
Are you able to test your own blood sugar level while in hospital?
Have you experienced any of the following difficult situations while in hospital (unexpectedly high blood sugar (hyper), unexpectedly low blood sugar (hopo), changes to meal times that affect control of your diabetes?, None of these)
Did staff respond appropriately to manage the situation?
Have you needed food to be brought into hospital in order to meet your dietary requirements and/or to manage your diabetes?
Has the hospital provided the right type of food for you to manage your diabetes
During this admission, how often was the choice of meal suitable for your diabetes?
During this admission, how often was the timing of meals suitable for your diabetes?
Do you feel that the hospital staff who look after you know enough about diabetes to meet your needs while in hospital
If you have had questions about your diabetes, were staff able to answer these in a way you could understand?
While in hospital, have you received enough emotional support from staff to manage your diabetes
How good do you think the staff are at working together as a team in managing your diabetes
How satisfied are you with the overall care of your diabetes while you were in hospital?
If there is anything else you would like to tell us about the diabetes care you have received during this hospital stay, please do so in here, e.g. things that could be improved or anything you found particularly good or bad about your care
NATIONAL DIABETES INPATIENT AUDIT 2013 (Bedside Audit Questionnaire, completed by medical staff)

SECTION A. BACKGROUND INFORMATION

Specialty of Ward, Speciality of consultant,

Patient age, Patient gender, Patient ethnic background

Patient diabetes type on admission (eg T1, T2 insulin treated, diet only, pancreatitis, MODY etc)

Is the patient being treated with sulponylurea, Is the patient having enteral feeding, How long has the patient had diabetes, Was the patient cognitively impaired at the time of the audit, Number of nights in hospital, type of admission (elective, emergency, transfer), Main reason for admittance (DKA, hyperglycaemia with established diabetes, active diabetic foot disease, non-medical etc),

KNOWN DIABETIC COMPLICATIONS

Receiving renal replacement therapy, Foot disease,

SECTION B. DIABETES CONTROL

Glucose chart available for review, Patient currently on intravenous insulin infusion? Looking at the previous 7 days, on how many days has blood glucose monitoring been carried out? On how many of these days was the frequency of monitoring appropriate?

On the days identified, i.e. in the previous 7 days, and counting only blood glucose readings separated by a 4 hr period: No. of glucose readings between 3-3.9 (or <3) mmol/L, Was the treatment of hypoglycaemia documented. Was the treatment in accordance with local guidelines,

No. of episodes of hypoglceamia requiring injectable treatment . If there has been hypoglycaemia during the last 7 days, please indicate the number of episodes in each of the following time periods etc

No. good diabetes days' in the last 7 days, defined as days in which the frequency of tests is appropriate (Q18) and there is no more than one reading > 11 mmol/L

What level of control is appropriate for this patient

Did the patient develop DKA at any time after their admission?

SECTION C. PRESCRIBING AND DRUG MANAGEMENT ERRORS OVER LAST 7 DAYS

Did the patient receive insulin at any time during the last 7 days? Was the drug chart available for review?

Did any of the following occur… Oral Hypoglycaemic Agent (OHA) prescription errors…. Insulin prescription errors: …. Insulin management errors:…. OHA management errors:…..

SECTION D. INTRAVENOUS INSULIN INFUSIONS

Has the patient been on an insulin infusion during the last 7 days?

Thinking of the most recent use of an insulin infusion, please complete the following:

Duration of insulin infusion…

Was the use of insulin infusion deemed appropriate (e.g. not eating or drinking, etc), Ws the duration of the infusion appropriate? If discontinued, has the transfer to sc insulin been managed appropriately, Total number of glucose readings in the last 24 hours on infusion, Total number of glucose readings > 11mmol/L in the last 24 hours on infusion?, Total number of glucose readings < 4mmol/L in the last 24 hours on infusion,

SECTION E. INVOLVEMENT OF THE SPECIALIST DIABETES TEAM

Was the patient previously aware that the diabetes team is available to provide support to inpatients with diabetes and advice to ward staff? Has the patient asked to be referred to the diabetes team, Is there documented evidence of the patient being seen by a member of the diabetes team? Does the patient wish the diabetes team to be involved in the management of their diabetes whilst in hospital? Should the patient have been referred to the diabetes team

SECTION F. GENERAL FOOT CARE

Was there any documentation of a diabetic foot risk (for ulceration) assessment in the FIRST 24 hours of admission? Was there any documentation of a diabetic foot risk (for ulceration) assessment AFTER the first 24 hours of admission? Was the patient admitted with active foot disease? Did a foot lesion (eg heel ulcer) arise during this admission? Was the patient seen by a member of the foot MDT within 24 hours? Has there been input from the foot MDT in the last 7 days

ASA, American Society of Anesthesiologists classification; BMI, Body mass index; T1, type 1; T2, Type 2; DKA, diabetic ketoacidosis; hr, hour; OHA, oral hypoglycaemic agent; MDT, multidisciplinary team; MODY, maturity-onset diabetes of the young

Reference List

(1) Gibbons E, Fitzpatrick R. A structured review of patient-reported outcome measures (PROMs) for diabetes. Report to the department of health, 2009 . Patient-reported Outcome Measurement Group 2009.

(2) 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.

(3) 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.

(4) 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

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

(6) Bech P, Moses R, Gomis R. The effect of prandial glucose regulation with repaglinide on treatment satisfaction, wellbeing and health status in patients with pharmacotherapy naive Type 2 diabetes: a placebo-controlled, multicentre study. Qual Life Res 2003 Jun;12(4):413-25.

(7) Redekop WK, Koopmanschap MA, Stolk RP, Rutten GE, Wolffenbuttel BH, Niessen LW. Health-related quality of life and treatment satisfaction in Dutch patients with type 2 diabetes. Diabetes Care 2002 Mar;25(3):458-63.

(8) Matza LS, Boye KS, Yurgin N, Brewster-Jordan J, Mannix S, Shorr JM, et al. Utilities and disutilities for type 2 diabetes treatment-related attributes. Qual Life Res 2007 Sep;16(7):1251-65.

(9) Lloyd A, Nafees B, Gavriel S, Rousculp MD, Boye KS, Ahmad A. Health utility values associated with diabetic retinopathy. Diabet Med 2008 May;25(5):618-24.

(10) Bharmal M, Thomas J, III. Comparing the EQ-5D and the SF-6D descriptive systems to assess their ceiling effects in the US general population. Value Health 2006 Jul;9(4):262-71.

(11) Currie CJ, Poole CD, Woehl A, Morgan CL, Cawley S, Rousculp MD, et al. The financial costs of healthcare treatment for people with Type 1 or Type 2 diabetes in the UK with particular reference to differing severity of peripheral neuropathy. Diabet Med 2007 Feb;24(2):187-94.

(12) Gore M, Brandenburg NA, Dukes E, Hoffman DL, Tai KS, Stacey B. Pain severity in diabetic peripheral neuropathy is associated with patient functioning, symptom levels of anxiety and depression, and sleep. J Pain Symptom Manage 2005 Oct;30(4):374-85.

(13) Hoffman DL, Sadosky A, Alvir J. Cross-national burden of painful diabetic peripheral neuropathy in Asia, Latin America, and the Middle East. Pain Pract 2008 Jan;9(1):35-42.

(14) Vernon MK, Brandenburg NA, Alvir JM, Griesing T, Revicki DA. Reliability, validity, and responsiveness of the daily sleep interference scale among diabetic peripheral neuropathy and postherpetic neuralgia patients. J Pain Symptom Manage 2008 Jul;36(1):54-68.

(15) Vexiau P, Mavros P, Krishnarajah G, Lyu R, Yin D. Hypoglycaemia in patients with type 2 diabetes treated with a combination of metformin and sulphonylurea therapy in France. Diabetes Obes Metab 2008 Jun;10 Suppl 1:16-24.

(16) Tolle T, Xu X, Sadosky AB. Painful diabetic neuropathy: a cross-sectional survey of health state impairment and treatment patterns. J Diabetes Complications 2006 Jan;20(1):26-33.

(17) Glasziou P, Alexander J, Beller E, Clarke P. Which health-related quality of life score? A comparison of alternative utility measures in patients with Type 2 diabetes in the ADVANCE trial. Health Qual Life Outcomes 2007;5:21.

(18) Johnson JA, Pickard AS. Comparison of the EQ-5D and SF-12 health surveys in a general population survey in Alberta, Canada. Med Care 2000 Jan;38(1):115-21.

(19) Johnson JA, Pickard AS, Conner-Spady B. Longitudinal changes in health-related quality of life of people with diabetes compared to those without chronic conditions. In: Badia X, editor. Proceeedings of the 17th EuroQoL Plenary Meeting, Pamplona, Spain, 28th - 29th September 2010. 2000. p. 193-206.

(20) Clarke P, Gray A, Holman R. Estimating utility values for health states of type 2 diabetic patients using the EQ-5D (UKPDS 62). Med Decis Making 2002 Jul;22(4):340-9.

(21) Wee HL, Tan CE, Goh SY, Li SC. Usefulness of the Audit of Diabetes-Dependent Quality-of-Life (ADDQoL) questionnaire in patients with diabetes in a multi-ethnic Asian country. Pharmacoeconomics 2006;24(7):673-82.

(22) Sakthong P, Charoenvisuthiwongs R, Shabunthom R. A comparison of EQ-5D index scores using the UK, US, and Japan preference weights in a Thai sample with type 2 diabetes. Health Qual Life Outcomes 2008;6:71.

(23) Matza LS, Yurgin N, Boye KS, Malley K, Shorr JM. Obese versus non-obese patients with type 2 diabetes: patient-reported outcomes and utility of weight change. Curr Med Res Opin 2007 Sep;23(9):2051-62.

(24) Matza LS, Boye KS, Yurgin N. Validation of two generic patient-reported outcome measures in patients with type 2 diabetes. Health Qual Life Outcomes 2007;5:47.

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

(26) 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.

(27) 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.

(28) Cummins E, Shyangdan D, Royle P, Johnston R, Ho V, Briscoe S, et al. Fluocinolone acetonide intravitreal implant for the treatment of diabetic macular oedema TA271. Warwick Evidence; 2012. Report No.: TA271.

(29) National Institute for Health and Care Excellence. Lucentis (ranibizumab) for the treatment of visual impairment due to diabetic macular oedema TA237. 2010. Report No.: TA237.

(30) Black C, Cummins E, Royle P, Philip S, Waugh N. The clinical effectiveness and cost effectiveness of inhaled insulin in diabetes mellitus: A systematic review and economic evaluation (TA113). 2006. Report No.: TA113.

(31) National Institute for Health and Care Excellence. Macular Oedema (diabetic) - Pegaptanib Sodium (ID452). 2011. Report No.: ID452.

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

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

(34) 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.

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

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

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

(38) Heintz E, Wirehn AB, Peebo BB, Rosenqvist U, Levin LA. QALY weights for diabetic retinopathy--a comparison of health state valuations with HUI-3, EQ-5D, EQ-VAS, and TTO. Value Health 2012 May;15(3):475-84.

(39) Eli Lilly and company. Exenatide Once Weekly (Bydureon) for the treatment of type 2 diabetes TA248. 2011. Report No.: TA248.

(40) Health and Social Care Information Centre. National Diabetes Audit (Adult) including Inpatient Audit (NADIA). Health and Social Care Information Centre 2014 [cited 2014 May 13];

(41) Bansback N, Czoski-Murray C, Carlton J, Lewis G, Hughes L, Espallargues M, et al. Determinants of health related quality of life and health state utility in patients with age related macular degeneration: the association of contrast sensitivity and visual acuity. Qual Life Res 2007 Apr;16(3):533-43.

(42) Czoski-Murray C, Carlton J, Brazier J, Young T, Papo NL, Kang HK. Valuing condition-specific health states using simulation contact lenses. Value Health 2009 Jul;12(5):793-9.

(43) Donnelly LA, Morris AD, Frier BM, Ellis JD, Donnan PT, Durrant R, et al. Frequency and predictors of hypoglycaemia in Type 1 and insulin-treated Type 2 diabetes: a population-based study. Diabet Med 2005 Jun;22(6):749-55.

(44) Levy AR, Christensen TL, Johnson JA. Utility values for symptomatic non-severe hypoglycaemia elicited from persons with and without diabetes in Canada and the United Kingdom. Health Qual Life Outcomes 2008;6:73.