
Article:
Gomersall, T., Spencer, S., Basarir, H. et al. (4 more authors) (2015) Measuring quality of life in children with speech and language difficulties: a systematic review of existing approaches. International Journal of Language and Communication Disorders, 50 (4). pp. 416-435. ISSN 1368-2822

https://doi.org/10.1111/1460-6984.12147

Measuring quality of life in children with speech and language difficulties: A systematic review of existing approaches

Tim Gomersall, PhD.,[1] Sarah Spencer, PhD.,[2] Hasan Basarir, PhD.,[1] Aki Tsuchiya, PhD.,[1,3] Judy Clegg, PhD.,[2] Anthea Sutton MCLIP.,[1] Kath Williams, MA.[1]

1: School of Health and Related Research, University of Sheffield
2: Human Communication Sciences, University of Sheffield
3: Department of Economics, University of Sheffield

Acknowledgments

We would like to thank Dr. Katherine Stevens and Dr. Jenny Owen for their helpful comments on a previous draft of this article. We are also grateful to the Faculty of Medicine, Dentistry, and Health, University of Sheffield, for providing us with funding via a Research and Innovation Award ("Paediatric speech and language impairment condition-specific outcome measure", PI Sarah Spencer) to make this work possible.

Abstract

Background: Children's speech and language difficulties (SaLD) can affect various domains of quality of life (QoL), and speech and language therapy interventions are an important way to address this. Systematically measuring QoL outcomes in this population is highly complex due to factors such as heterogeneity in impairments and differing therapeutic targets. However, measurements of QoL are increasingly demanded by health care commissioners and policy-makers to inform resource allocation.

Objectives: To review the use of QoL outcome measures in research involving children (age ≤18 years) with speech and language difficulties

Methods: Systematic review. A systematic search across various databases was performed. Information on methodological details of each relevant study, along with descriptions of the QoL measures employed, were extracted into standardised data extraction forms. Findings were discussed in a narrative synthesis.

Findings: Nineteen relevant studies were identified, dealing with a range of sub-populations of children with SaLD. For the most part, generic QoL measures were used, although there was little convergence on the type of QoL measures employed throughout the literature. Five studies utilised preference-based QoL measures, including the 16D / 17D, HUI3, EQ-5D, and QWB-SA. Of these measures, the HUI3 demonstrated the most promising discriminant validity, although the preference weights for this measure were generated with adults.

Conclusions: Quality of life outcomes among children with SaLD are not yet being captured in a systematic way. The HUI3 measure appears to show some promise for generating preference-based health utilities in this population, although further testing of the measure is required.

Key words: Speech and language difficulties; paediatric; quality of life; outcomes; systematic literature review

Measuring quality of life in children with speech and language difficulties: A systematic review of existing approaches

Introduction

This article explores quality of life measurement in children and adolescents with speech and language difficulties (SaLD). Specifically, we sought to overview the generic and condition-specific health-related quality of life (HRQoL) measures currently being used in the field, and whether any such measures were suitable for informing service delivery priorities within this population. Because health care needs among children with SaLD tend to be heterogeneous, with the focus being the child's specific needs (Pring, Flood, & Dodd, 2012), it has proven difficult to apply outcome measures in a consistent and systematic way across the population of children with SaLD. Nevertheless, recent evidence from qualitative research suggests that children with SaLD and their families value some common outcomes from the care they receive pertaining to quality of life – such as improved social participation, confidence and independence (Markham, van, Gibbard, & Dean, 2009; Markham & Dean, 2006).

Children's SaLD result in poor communication and quality of life for children and their families, and have associated long-term implications for psychosocial outcomes such as future employability (Clegg, Hollis, Mawhood, & Rutter, 2005), academic attainment (Conti-Ramsden, Botting, Simkin, & Knox, 2001; Clegg et al., 2005), and mental health (Conti-Ramsclen, Batting, & Durkin, 2008). However, little is known about the measurement of HRQoL in this population, and the extent to which interventions may affect this. A recent systematic review of the health related quality of life of children with SaLD found that seven studies had been conducted in this area, with all reporting that SaLD had a substantial impact on children's quality of life (Feeney, Desha, Ziviani, & Nicholson, 2012). Another recent review of emerging evidence has found that SLT interventions for children's SaLD can be effective during research trials, though less is known about effectiveness during clinical implementation (Law et al., 2010). It is also less clear whether systematic measurement of outcomes has been adequately tailored to capture the particular issues encountered in SaLD interventions and support, since these are complex in terms of:

a) Systems involved (families, schools, multi-disciplinary health care teams);

b) Multiple intervention components (e.g. speech sound therapy, strategies for participating in successful communication, training for families);

c) Heterogeneity of expected outcomes (e.g. intelligibility, improved language understanding, facilitated access to education); and

(d) Time horizon of the analysis (e.g. direct measurement limited to the duration of the trial, longer term modelling, modelling of other unobserved outcomes).

Measuring outcomes is further complicated by the level targeted by intervention, which may concern different outcomes for an individual child, for a class or a school, for a speech and language therapy

service or indeed for the population of a local authority. Furthermore, given the heterogeneity of impairments associated with SaLD, the definition of a 'successful' intervention differs for individual children according to not only the nature of the outcome but also in the magnitude of change following intervention.

Despite the challenges of measuring outcomes of intervention for paediatric SaLD, establishing an appropriate outcome measure is a priority. Many SLT interventions are provided via healthcare services, where there is a long-recognised need to develop methods for reporting the impact of health services (Wilkin, Hallam, & Doggett, 1992). Furthermore, given the shift towards Payment by Results and the emphasis on outcomes-based procurement within the National Health Service (NHS) in the UK, there is increased pressure to demonstrate benefits of interventions (Boyle, 2011; Enderby, John, & Petherham, 2006). This is an issue for the SLT profession with respect to outcome measurement; and it is imperative that any measures used are sufficiently sensitive to demonstrate small but important changes for children with SaLD, and specifically attuned to the particular ways in which SaLD impact the QoL of children and their families.

There are roughly two types of motivation for the demand for outcome measures. First, a descriptive system of relevant outcomes will facilitate the comparison across patients, or the monitoring of a given patient over time. Typically, a descriptive system will be made up of a number of items (or questions) each with two response categories (e.g. 'yes' or 'no'), or more levels (e.g. 'always', 'often', 'sometimes', or 'never'). Different items can be grouped into dimensions (or attributes). Responses to the items could be elicited from the patient, a carer, or a clinician. Descriptive systems that can be used across different condition are called generic (e.g. SF-36, (Ware & Sherbourne, 1992); otherwise, they are condition-specific (e.g. GHQ-12 for mental health, (Goldberg, 1972)).

Outcome measures can also provide an assessment of the states. In order to accommodate this motivation, most descriptive systems provide summary scores by adding up the responses to individual items. However, this implies that, for instance, the effect of moving up a notch in two different items is equivalent to moving up two notches in any other single item. Hence, unless it is assumed that all levels across all items are equally important, a simple sum of item scores only provides a very approximate picture. An outcome measure that allows for the different levels across the items to have different value is called a preference-based outcome measure. Such a measure is made up of a descriptive system and an accompanying value set that reflects the relative importance of the levels and items, obtained through a valuation study.

This second motivation is linked to the need in resource allocation decisions at the policy level to capture improvements in HRQoL on a scale that is commensurable with improvements in survival. To achieve this, the preference weights, or utility scores, are anchored at 1 for full health and 0 for

being dead. Thus, living in full health for a year is regarded as equivalent to living in another health state valued at 0.5 for 2 years, both of them representing one Quality Adjusted Life Year (QALY). In this particular context, it becomes important that, wherever possible, estimation of QALYs gained is based on generic preference-based outcome measures such as the EQ-5D (Brooks & the EuroQol Group, 1996) or SF-6D (Brazier, Roberts, & Deverill, 2002) so that the outcomes are measured on a common comparable scale. Various interventions for different disease areas are competing for funds from the same health budget and therefore, are required to be comparable in terms of mortality and morbidity by using a 'common currency' (Whitehead & Ali, 2010). However, generic outcome measures are designed to cover broad brush attributes of health, and thus may miss important aspects of health that are unique to specific conditions. Therefore, where generic measures are judged to be inappropriate, condition-specific preference-based measures (e.g. AQL-5D (Yang, Brazier, Tsuchiya, & Young, 2011) for asthma; or EORTC-8D (Rowen et al., 2011) for cancer) have been developed.

This article investigates the availability and use of HRQoL outcome measures in relation to children's SaLD, and particularly whether preference-based measures have been used in this context. It aims to investigate whether the existing outcome measures are fit for purposes of describing and assessing paediatric SaLD states.

Research questions:

1. What generic and condition-specific HRQoL outcome measures are used within the field of paediatric speech and language therapy, if any?

2. Have any outcome measures used within the field been:

    a. Validated for use with paediatric SaLD?

    b. Applied to evaluating the impact of SLT or psychosocial intervention?

    c. Weighted according to population preference, so that QALY losses or gains may be calculated?

Methods

Identification of studies

Extensive searches were undertaken with the aim of a comprehensive retrieval of studies relating to the research questions.

The search strategy comprised the following main elements:

- searching of electronic databases listed below
- scrutiny of bibliographies of retrieved papers and previous systematic reviews

Sources searched

The following electronic databases were searched from inception: Medline, Web of Science, ASSIA, ERIC, EMBASE, the Cochrane Databases, AMED, CINAHL, and PsycINFO.

Search terms

The search strategy was developed in collaboration with information specialists. Search terms included 'pediatrics', 'adolescents', and a broad variety of related terms (e.g. 'schoolchild*', 'teen*', 'youth*'), which were combined with SaLD-relevant terms (e.g. 'communication disorders', 'language delay*'), and terms pertaining to HRQoL specifically and outcome measurement more generally, in order to scope the literature broadly. Searches were aggregated with the Boolean operators AND / OR. A sample search strategy from Web of Science is included in Appendix 1. The last date of the searches was January 2013

Inclusion criteria

Population

The population comprised children and adolescents (age $\leq$ 18 years) with any speech and/or language impairment. Studies which also included participants with non-SaLD disabilities were included if data relating to participants with speech and/or language impairments were available separately, or if the proportion of participants with non-SaLD disabilities was small (<10%).

Intervention(s)

Any type of intervention was included in the review. However, we also included any scale development and/or validation studies if these were explicitly targeted for children with SaLD. Studies of surgical interventions in potentially relevant populations (e.g. palatoplasty for cleft palate) were not included.

Evaluation

Only studies which used at least one QoL measure in relation to the population of interest were eligible for inclusion.

Study selection

Using the above inclusion criteria, studies were screened through a two-stage process. The references collated in the Reference Manager© database were assessed for relevance first by title/abstract, and then by full text, excluding at each step studies which did not meet our criteria. A systematic reviewer (TG) examined titles and abstracts for inclusion, and screening was checked by a second reviewer with SLT expertise (SS) on a random sample of 50 citations. The kappa coefficient was calculated to measure inter-rater reliability. This was very high, at 1.0, indicating no discrepancies.

Data extraction strategy

Details of study design, population, sample size, and other aspects of methodology were inputted into a standardised data extraction form (appendix 2). Details of the QoL measures were also extracted, specifically 1. The name of the measure; 2. The dimensions described by the measure; 3. The specific items listed on the measure; 4. Scoring method; and 5. Whether the measure had been tested for validity and/or reliability with the target population; and if so, with what results. Where information on scale validity or reliability was not presented adequately in the included articles, further details were sought by searching Medline and/or PSYCInfo for previous validation studies of the measures. The data described above were then summarised in tabular form, and discussed in a narrative synthesis.

Results

Study characteristics

The electronic literature searches identified 5625 citations, 5603 of which were excluded at the title or abstract stage, leaving 23 which were obtained for examination of the full text. Six further articles were excluded at full-text stage, leaving a sample of 19 articles representing 19 studies. (Figure 1).

FIGURE 1 ABOUT HERE

Most studies (n=17) had a cross-sectional design. Also identified was a prospective cohort study (McDougall, Wright, Miller, & Nichols, 2012), and a scale development study (Thomas-Stonell, Oddson, Robertson, & Rosenbaum, 2010). The most commonly examined populations were children and/or adolescents with autism/asperger syndrome (Kuhlthau et al., 2010; Lee, Harrington, Louie, & Newschaffer, 2008; McDougall et al., 2012; Thomas-Stonell et al., 2010; Tilford et al., 2012), cerebral palsy (Majnemer, Shevell, Law, Poulin, & Rosenbaum, 2008; Petrou & Kupek, 2009; Thomas-Stonell et al., 2010; Keskin, Gunel, & Aktan, 2012), and specific language impairment (Arkkila et al., 2009; Arkkila et al., 2011; Flapper & Van den Heuvel, 2011; Flapper & Schoemaker, 2013). A number of studies described the QoL of children with SaLD alongside other chronic conditions (Lee et al., 2008; McDougall et al., 2012) (Willems et al., 2009). Sample size ranged from 22 (Limbers, Heffer, & Varni, 2009) to 6802 (Lee et al., 2008), although most of the sample in the latter study (n=6319) were children with a non-SaLD diagnosis (attention deficit hyperactivity disorder). The mean sample age (where reported) varied from 3.8 years (Thomas-Stonell et al., 2010) to 14.2 years (Arkkila et al., 2009). The research body covered an international base including the US and/or Canada (10 studies), Netherlands (3 studies), Finland (2 studies), the UK (1 study) and Europe-wide (1 study). This information is summarised in Table 1.

TABLE 1 ABOUT HERE

Details of included QoL measures

Thirteen measures were identified (Table 2), suggesting little convergence on any 'gold standard' QoL measure for use in SaLD research. The most commonly used measures were the paediatric generic PedsQL (Damiano et al., 2007; Keskin et al., 2012; Majnemer et al., 2008) and TACQOL(Flapper & Van den Heuvel, 2011; Flapper & Schoemaker, 2013; Willems et al., 2009). However, a variety of other measures were used in other studies, including the Kidscreen,(White-Koning et al., 2007) FOCUS(Thomas-Stonell et al., 2010) and EQ-5D(Willems et al., 2009). Five adult generic preference-basedmeasures had been used within the research (16D, 17D, HUI3, QWB-SA, and EQ-5D), although no condition-specific preference-based measures were seen. The number of individual items in each measure ranged from five (EQ-5D) to 63 (TACQOL), and the number of scale dimensions also varied substantially between studies. A factor analysis of the FOCUS instrument suggested it was capturing a single dimension – "real world communication outcomes", which was correlated with overall scores on the PedsQL measure (r=0.466, 95% CI: 0.053, 0.884, p=0.029). Conversely, the Kidscreen and NSCH measures were reported to measure 10 domains each.

TABLE 2 ABOUT HERE

Characteristics of measured domains and individual scale items

There was variation in the number or type of QoL domains measured with the different scales. Generic QoL scales such as the PedsQL; TACQOL/TAPQOL; KINDL-R; and EQ-5D, placed varying emphasis on SaLD-related domains, and often included domains of arguably tangential importance in SaLD (e.g. mobility and physical functioning, or pain). However, this may depend on whether the SaLD is a child's primary difficulty, or whether the SaLD exists with other impairments as part of one condition, for example in the case of mobility in cerebral palsy. Furthermore, it seems likely that many apparently generic QoL domains would be of relevance to QoL measurement in children with SaLD. The PedsQL, for instance, includes "emotional", "social", and "school" domains – and it is well known that SaLD can directly affect participation in such social activities. Indeed, condition-specific scales also examined similar issues. The ASHA Quality of Communication Life scale (QCL), for example, included items on emotional well-being ("I see the funny things in life"), social functioning (e.g. "Peers include me in conversations"), and school functioning (e.g. "I meet the communication needs of my job or school").

Validity and reliability

A number of studies reported previous or concurrent validation of the QoL scales they had used. Reliability of the PedsQL has previously been demonstrated in a number of studies, with internal consistency being good for total Scale Score (α= 0.88 child, 0.90 parent report), Physical Health Summary Score (α= 0.80 child, 0.88 parent), and Psychosocial Health Summary Score (α = 0.83

child, 0.86 parent). The PedsQL has been shown to discriminate between healthy children and those with chronic conditions, and to relate to indicators of illness burden (Varni, Seid, & Kurtin, 2001). It has also demonstrated sensitivity and responsiveness in various populations(Seid, Varni, & Kurtin, 2000; Varni et al., 2001; Varni, Burwinkle, Seid, & Skarr, 2003), although validation studies are yet to be reported in children with SaLD.

Initial testing of the TACQOL (n=77 non-SaLD outpatients from a paediatric clinic), revealed a range of Cronbach's α for the subscales from 0.71 to 0.89, and all but two of the individual items correlated more highly with their own subscales than with others (Vogels et al., 1998). The TACQOL's sister scale for younger children, the TAPQOL, was tested on preterm children (n=121) as well as general population (n=362). Cronbach's α ranged from 0.66 to 0.88 for the preterm children, and 0.43 to 0.84 for the comparison sample. The scale was shown to discriminate healthy from non-healthy children, and the a priori structure of the separate scales was demonstrated with principal components analysis, with low correlations between subscales (Fekkes et al., 2000). However, neither the TACQOL nor the TAPQOL appear to have been tested for validity or reliability in a sample of children with SaLD. The generic KINDL-R questionnaire has been widely tested, with Cronbach's α being >.70 for most subscales and samples. Psychometric testing has also revealed a good ability to distinguish between children with different types of physical chronic disorders and experiencing different types of strain (Bullinger, Brutt, Erhart, & Ravens-Sieberer, 2008). The properties of the scale have also been tested in children with stroke, and it was shown to discriminate stroke survivors from healthy controls, as well as neonatal vs childhood stroke survivors (Neuner et al., 2011). Both the SLSS and BMLSS, which were used in the McDougall study (McDougall et al., 2012), have demonstrated acceptable alpha coefficients (Bussing et al., 2009). Furthermore, the McDougall study indicated acceptable internal consistency for these scales in a sample of patients including those with autism, cerebral palsy, and general communication disorders (SLSS: $r = 0.79$ and BMLSS: $r = 0.72$).

Two measures (used in two studies) were designed specifically with people with SaLD: FOCUS and ASHA-QCL. Burgess et al. reported that the ASHA-QCL scale had previously undergone pilot testing with 57 adults with neurogenic communication disorders, while 86 adults with similar disorders participated in field testing. The mean level of inter-rater reliability on this testing was 0.63, and the scores were related to a general measure of well-being, the Affect Balance Scale(Bradburn, 1969). Although we were unable to identify any evidence of the reliability and validity of the ASHA-QCL with paediatric patients, Burgess(Burgess & Turkstra, 2010) reported that, when the rankings of items by children with autism and their parents were explored in a Spearman rank-order correlation, they were found to be linked ($r_s(16) = .54$, $p < 0.05$). The FOCUS was tested in a cohort of young children (<6 years of age) with any kind of communication impairment in Canada (n=165). When measured

against PedsQL scores, there was a correlation between higher overall PedsQL and FOCUS scores (r=0.466, p=0.029, CI: 0.053, 0.884). The FOCUS correlated especially well with psychosocial domain of PedsQL r=0.528, p=0.013, CI: 0.110, 0.842), and demonstrated good internal consistency for both parent (α=0.96) and clinician completers (α=0.94). However, since the scale has been designed for children under six (and for proxy completion), it is unknown whether a modified and/or self-report version of the scale may be required for older children and/or those with less severe impairments.

Studies that utilised preference-based QoL measures

Five studies used five preference-based measures; they were: (Arkkila et al., 2009) (Arkkila et al., 2011; Petrou & Kupek, 2009; Tilford et al., 2012). The Petrou and Kupeck study drew on data from a nationwide, UK-based survey ("Disability Survey 2000: Survey of Young People With a Disability and Sport") to estimate health utilities associated with a large range of chronic childhood conditions. As part of this survey, parents of children aged 5 to 16 years were asked to complete the Health Utilities Index (HUI3); a generic preference-based QoL instrument for adults that measures eight domains of HRQoL, with each domain having five or six functional levels (Table 2). Disutilities relative to population norms were presented for a number of relevant conditions: ASD: -0.494 (95% CI: -0.372, -0.624); Down's syndrome: -0.566 (95% CI: -0.450, -0.691); cerebral palsy: -0.652 (95% CI: 0.536, 0.775); and speech disorders: -0.487 (95% CI: -0.314, -0.668).

Willems (Willems et al., 2009) explored the use of EQ-5D as a measure for children with various chronic conditions, including a subset of children with speech, language, and / or hearing difficulties. EQ-5D is a generic preference-based HRQoL measure for adults, assessing five domains of well-being (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) which is widely used in health economic evaluations. The study found low to moderate correlations between the EQ-5D and generic TACQOL, with the highest correlation being between EQ-5D utility scores and the 'autonomy' dimension of the TACQOL(r=.57). In contrast to TACQOL, however, EQ-5D was shown to have poor discriminant power, particularly with respect to identifying those with SaLD compared to other conditions.

Tilford (Tilford et al., 2012), compared the sensitivity of two generic preference-based HRQOL instruments for ASD-related symptoms and conditions: the HUI 3 and the Quality of Well-Being Self-Administered (QWB-SA). The QWB-SA is a self-administered preference-based generic measure for adults in which a composite score of 58 symptoms and problems is combined with three scales of functioning (mobility, physical activity and social activity) to provide a health utility score ranging from 0 (being dead) to 1.0 (asymptomatic and fully functioning). Both the QWB-SA and HUI3 were

examined for the domains contributing most to health utility scores, and were correlated with ASD-specific diagnostic instruments to assess their validity. For both instruments, speech problems represented the highest percentage of problem responses: on the HUI 3, 20.7% of caregivers reported that their child had level 1 speech (able to be understood completely when talking with strangers or friends), with 78.7% reporting at least some level of speech impairment. Similarly, on the QWB-SA, 51.4% of caregivers reported that their child had been stuttering / unable to speak clearly over the past 3-day period.

In terms of ability to discriminate different ASD diagnoses, the HUI3 demonstrated a greater disutility associated with autistic disorder (0.64 $\pm$0.23, range: 0.07, 1.0) than pervasive developmental disorder-not otherwise specified (0.70 $\pm$0.24, range: -0.03, 0.93), which in turn showed greater disutility than Asperger's syndrome (0.79 $\pm$ 0.16, range: 0.57, 1.0). However, while the QWB-SA was able to show greater disutility associated with autistic disorder (0.58 $\pm$ 0.16, range: 0.18, 1.0) than pervasive developmental disorder-not otherwise specified (0.62 $\pm$ 0.18, range: 0.27, 1.0), a significant difference was not seen between the latter diagnosis and Asperger's syndrome (0.62 $\pm$ 0.15, range: 0.36, 0.89). With respect to composite scores on a commonly used measure of autism severity, the Vineland II adaptive behaviour scales, the HUI3 demonstrated a moderate, significant correlation (ρ =0.521, p<0.001), while the QWB-SA showed a weaker, though still statistically significant correlation (ρ =0.247, p<0.001). There was a mixed picture with respect to correlations with caregiver-reported symptoms; however, the HUI3 scores were consistently better correlated than the QWB-SA scores with clinician-rated symptoms. Consequently, Tilford et al (Tilford et al., 2012) suggest that, while there was indication of construct validity for both scales, the HUI3 is likely to be a better instrument for estimating any quality of life gains among children with ASD, and is also better placed to capture health states associated with language impairment.

Arkkila et al. performed two studies (Arkkila et al., 2009) (Arkkila et al., 2011) to explore the viability of the 16D and 17D, respectively, for comparing the HRQoL of adolescents and children with specific language impairment (SLI) with typically developing peers. In the first of these studies(Arkkila et al., 2009), the 16D was mailed to adolescents with SLI, with responses being obtained for 49 participants (73% response rate). Responses were compared with an age- and gender-matched sample drawn from a national survey in Finland. There was little difference between the SLI group and controls on the overall 16D health utility scores (SLI group: 0.946, controls: 0.951). However, the SLI group showed a significantly lower score on the domain of mental function (p=0.001), while the control group scored lower on vitality (p=0.003). Interestingly, although there was a difference in the speech domain, this failed to reach statistical significance once the p-value

threshold was adjusted for multiple comparisons (p=0.023). In the next study (Arkkila et al., 2011), a sample of children aged 8 to 11 years were compared to a population sample using 17D. Again, there was no significant difference between the SLI and control samples on overall scores (SLI: 0.9337; control: 0.9333). Significant differences in the responses of children with SLI to population controls were observed for the domains of sleep and speech. In terms of sleep, the general population appeared to do worse than children with SLI: 65% of the SLI group were rated as level 1 (best possible), and the remaining 35% were rated as level 2. Conversely, 42% of the general population were rated as level 1; 52% as level 2; 4% as level 3; and 2% as level 4 (p=0.001 for this comparison). In the domain of speech, the percentage of children with SLI rated as levels 1, 2, 3, and 4 were 63%, 33%, 2%, and 2%, respectively. The population sample values for comparison were 96%, 3%, 0%, and 0.8%, respectively (p<0.001). In addition, the 17D correlated poorly with a number of relevant diagnostic tests carried out as part of the study (including the Wechsler Intelligence Scale for Children-Revised, and the Neuropsychological Test Battery for Children).

Discussion

This review examined the measurement of HRQoL in children with SaLD, particularly which instruments are currently being used in research, whether these instruments have been validated, and whether any of the measures were preference-based so that could provide data for NICE economic evaluations in terms of cost/QALY. It was found that some efforts have recently been underway to estimate the HRQoL of children with SaLD, with a small but growing corpus of literature providing data on QoL decrements associated with SaLD. Generic measures, such as the PedsQL and TACQOL were in fairly common use among the literature, while condition-specific measures were less frequently employed. Indeed, in a surprising finding of the review, no studies using Enderby's condition-specific Therapy Outcome Measures (TOM) approach (Enderby et al., 2006) were identified. It was unclear why this apparently slow uptake of TOM to describe the quality of life in the paediatric setting was observed, although TOM does appear to have been more widely used in older patients, especially stroke survivors. There is an argument for the use of generic preference-based HRQoL instruments, in that they are preferred by policymakers, and provide a simple way to examine outcomes across different conditions on a comparable scale. Yet generic instruments are likely to miss some important dimensions of psychosocial functioning and well-being specific to children with SaLD. In the UK context, in which NICE uses the QALY measure to inform healthcare resource allocation, measures used for this purpose need to be sensitive and specific enough to pick up any small but important differences of relevance to children with SaLD and their carers. Rather than generic measures, condition-specific outcome measures may be required to fairly assess outcomes in paediatric SaLD research. However, existing condition-specific measures are not preference-based and therefore cannot be used to capture the benefits of interventions in terms of QALYs.

It was therefore clear that much work is still required for robust implementation of HRQoL measures in general, and preference-based measures in particular, among the paediatric SaLD population. Although the validity and/or reliability of most of the scales had been assessed in previous work, this was seldom with children/adolescents with SaLD. Consequently, the reliability and/or validity of many existing generic measures with respect to capturing HRQoL changes in children with SaLD remains uncertain. A further potential threat to the reliability of HRQoL measurement was the widespread use of proxy reports throughout the literature. Arguably, proxies are mandated by both the limitations of young children's capacity to describe how a health condition is affecting their HRQoL, and the fact that the children's capacity to do so is further compromised by the presence of SaLD. Proxy reports are an important way to gather data on the perspectives of those who have problems expressing themselves verbally, and indeed, for a substantial subset of children with SaLD, self-reported HRQoL measurement may simply not be possible. Nevertheless, it cannot be assumed that a proxy report provides an accurate picture of HRQoL from the perspective of a child with SaLD, and so caution should be exercised when using this approach. Two of the articles included in this review (Majnemer et al., 2008; White-Koning et al., 2007) specifically set out to investigate the agreement in QoL ratings between children with SaLD and their carers, and both these studies found some discrepancies in child and carer ratings. However, measures designed for adults with communication impairments include items that may be difficult for children with SaLD to comprehend. For example, items such as "I meet the communication needs of my job or school" in the QCL scale are couched in rather abstract terms, and thus may require rephrasing to be viable for use with children with SaLD. Notwithstanding the general difficulties in administering survey measures to people with communication needs, studies such as Assessing Communication Therapy in the North West of England (Young, Gomersall, & Bowen, 2013) have shown that it may be possible to access the perspectives of adults with even quite severe acquired communication impairments, provided that appropriate support is provided. Similar techniques, for example using visual prompts rather than standard discursive approaches, are also showing some promise in childhood research (for a discussion, see (James, 2007)). However, the potential role for such methods in child-centred instrument design is yet to be fully explored.

With respect to preference-based measurement of HRQoL in children with SaLD, there was a small but growing level of research in the area, with five studies exploring five different measures. It appears that the HUI3 may be the best option at present for assessing the QoL of children with SaLD. The measure was reliably able to demonstrate disutilities associated with a range of SaLD conditions in children and adolescents, and was sensitive to different ASD diagnoses. The HUI3 also showed moderate correlations with measures of autism symptom severity, especially the Vineland Adaptive Behavior Scales and clinician-rated symptoms. While these are promising signs for the HUI3, it cannot be assumed that the findings from studies of children with ASD generalise to the wider SaLD

population. Indeed, the measure included QoL constructs that were of limited relevance to the population of interest (e.g. ambulation, dexterity, pain), and it is far from certain that the measure would pick up small but important differences in a programme evaluation. Willems et al. (Willems et al., 2009), by contrast, showed that the generic EQ-5D had a poor ability to discriminate children with SaLD from the general population, and was better placed to assess the health utilities of children with rheumatic disorders. Arkkila et al. (Arkkila et al., 2009; Arkkila et al., 2011) reported mixed findings on the 16D and17D measures, respectively, with both measures performing poorly on discriminating children with SLI from the general population on overall health utility scores, yet also showing some ability to discriminate children with SLI on some domains. Furthermore, no single study has compared three or more generic preference-based instruments alongside each other.

As with any piece of secondary research, this review has a number of limitations. Data were taken from various types of study, and a heterogeneous set of conditions fall under the rubric of paediatric SaLD. Hence, the use of a measure with one subset of children with SaLD does not necessarily mean that the appropriateness of the measure is generalizable to the wider paediatric SaLD population, or to different research contexts. In addition, although a thorough search strategy was employed, it is possible that some literature was missed. Finally, it should be noted that we have only focused on the specifics of outcome measurement in this review. In terms of the broader issue of evaluating complex interventions such as SLT, a multiplicity of methods is called for, and it was beyond the scope of this article to consider other approaches in any detail. For example, qualitative research may be useful to develop condition-specific measures reflecting patient priorities, to explore the perceived reasons for programme (in)effectiveness, and to investigate how effective interventions may be transferred outside the controlled context of clinical trials (Campbell et al., 2000; Campbell et al., 2007).

In conclusion, the HUI3 appears to be the most promising generic, preference-based HRQoL measure for use in paediatric SaLD research. However, as a generic measure, it is also possible that the HUI3 may not be adequately attuned to SaLD-related issues to capture small but important changes after intervention with this group. Further research is warranted into the sensitivity of HUI3 to such changes, and it may be that 'bolt-on' HRQoL dimensions would be needed to make existing generic measures more relevant to HRQoL measurement in children. Furthermore, the HUI3 preference weighting system was developed for adults, and so the applicability of these weights to children is associated with a great deal of uncertainty. In the political and economic context of the contemporary UK healthcare system, in which funding is prioritised for those interventions which can demonstrate meaningful changes in patients' health, ideally in terms of QALYs gained, it is vital that SLT practitioners use sensitive, appropriate measures to demonstrate the benefits of their interventions for children with SaLD and their families.

Reference List

Arkkila, E., Rasanen, P., Roine, R. P., Sintonen, H., Saar, V., & Vilkman, E. (2009). Health-related quality of life of adolescents with childhood diagnosis of specific language impairment. International Journal of Pediatric Otorhinolaryngology, 73, 1288-1296.

Arkkila, E., Rasanen, P., Roine, R. P., Sintonen, H., Saar, V., & Vilkman, E. (2011). Health-related quality of life of children with specific language impairment aged 8-11. Folia Phoniatrica et Logopedica, 63, 27-35.

Bercow, J. (2008). The Bercow Report: A Review of Services for Children and Young People (0–19) with Speech, Language and Communication Needs .

Ref Type: Bill/Resolution

Bishop, A. G. & League, M. B. (2006). Identifying a Multivariate Screening Model to Predict Reading Difficulties at the Onset of Kindergarten: A Longitudinal Analysis. Learning Disability Quarterly, 29, 235-252.

Boyle, S. (2011). United Kingdom (England): Health system review. Health Systems in Transition, 13, 483.

Bradburn, N. M. (1969). The structure of psychological well-being. Chicago: Aldine.

Brazier, J., Roberts, J., & Deverill, M. (2002). The estimation of a preference-based measure of health from the SF-36. J.Health Econ., 21, 271-292.

Brooks, R. & the EuroQol Group (1996). EuroQol: The current state of play. Health Policy, 37, 53-72.

Bullinger, M., Brutt, A. L., Erhart, M., & Ravens-Sieberer, U. (2008). Psychometric properties of the KINDL-R questionnaire: results of the BELLA study. European Child and Adolescent Psychiatry, 17 Suppl 1, 125-132.

Burgess, S. & Turkstra, L. S. (2010). Quality of communication life in adolescents with high-functioning autism and Asperger syndrome: a feasibility study. Language, Speech & Hearing Services in the Schools, 41, 474-487.

Bussing, A., Fischer, J., Haller, A., Heusser, P., Ostermann, T., & Matthiessen, P. F. (2009). Validation of the brief multidimensional life satisfaction scale in patients with chronic diseases. Eur.J.Med.Res, 14, 171-177.

Campbell, M., Fitzpatrick, R., Haines, A., Kinmonth, A. L., Sandercock, P., Spiegelhalter, D. et al. (2000). Framework for design and evaluation of complex interventions to improve health. BMJ, 321, 694-696.

Campbell, N. C., Murray, E., Darbyshire, J., Emery, J., Farmer, A., Griffiths, F. et al. (2007). Designing and evaluating complex interventions to improve health care. BMJ, 334, 455-459.

Clegg, J., Hollis, C., Mawhood, L., & Rutter, M. (2005). Developmental language disorders--a follow-up in later adult life. Cognitive, language and psychosocial outcomes. Journal of Child Psychology & Psychiatry & Allied Disciplines, 46, 128-149.

Conti-Ramsclen, G., Batting, N., & Durkin, K. (2008). Parental perspectives during the transition to adulthood of adolescents with a history of specific language impairment (SLI). Journal of Speech Language and Hearing Research, 51, 84-96.

Conti-Ramsden, G., Botting, N., Simkin, Z., & Knox, E. (2001). Follow-up of children attending infant language units: outcomes at 11 years of age. International Journal of Language & Communication Disorders, 36, 207-219.

Damiano, P. C., Tyler, M. C., Romitti, P. A., Momany, E. T., Jones, M. P., Canady, J. W. et al. (2007). Health-related quality of life among preadolescent children with oral clefts: The mother's perspective. Pediatrics, 120, E283-E290.

Enderby, P., John, A., & Petherham, B. (2006). Therapy outcome measures for rehabilitation professionals. Chichester: John Wiley & Sons, Ltd.

Feeney, R., Desha, L., Ziviani, J., & Nicholson, J. M. (2012). Health-related quality-of-life of children with speech and language difficulties: a review of the literature. [Review]. International Journal of Speechlanguage Pathology, 14, 59-72.

Fekkes, M., Theunissen, N. C., Brugman, E., Veen, S., Verrips, E. G., Koopman, H. M. et al. (2000). Development and psychometric evaluation of the TAPQOL: a health-related quality of life instrument for 1-5-year-old children. Qual.Life Res, 9, 961-972.

Flapper, B. C. & Schoemaker, M. M. (2013). Developmental Coordination Disorder in children with specific language impairment: Co-morbidity and impact on quality of life. Research in Developmental Disabilities, 34, 756-763.

Flapper, B. C. & Van den Heuvel, M. (2011). Health-related quality of life in school-age children with speech-language-impairment. Journal of Neurology, 258, 233-234.

Frank, S. M. & Rieber, R. W. (1981). Language development and language disorders in children and adolescents. In R.W.Rieber (Ed.), Communication disorders (pp. 179-205). New York: Plenum Press.

Goldberg, D. P. (1972). The detection of psychiatric illness by questionnaire. London: Oxford University Press.

James, A. (2007). Giving voice to children's voices: Practices and problems, pitfalls and potentials. American Anthropologist, 109, 261-272.

Keskin, N., Gunel, M. K., & Aktan, T. (2012). Health related quality of life of adolescents with cerebral palsy: Self versus parent perspectives. Developmental Medicine and Child Neurology [On-line].

Kuhlthau, K., Orlich, F., Hall, T. A., Sikora, D., Kovacs, E. A., Delahaye, J. et al. (2010). Health-Related Quality of Life in children with autism spectrum disorders: results from the autism treatment network. Journal of Autism and Developmental Disorders, 40, 721-729.

Law, J., Garrett, Z., & Nye, C. (2010). **Speech and language therapy interventions for children with primary speech and language delay or disorder**. The Cochrane Library 5.

Ref Type: Journal (Full)

Lee, L. C., Harrington, R. A., Louie, B. B., & Newschaffer, C. J. (2008). Children with autism: quality of life and parental concerns. Journal of Autism & Developmental Disorders, 38, 1147-1160.

Limbers, C. A., Heffer, R. W., & Varni, J. W. (2009). Health-related quality of life and cognitive functioning from the perspective of parents of school-aged children with Asperger's Syndrome utilizing the PedsQL. Journal of Autism and Developmental Disorders, 39, 1529-1541.

Lindsay, G., Dockerall, J., Law, J., oulstone, S., & Vignoles, A. (2011). Better Communication Research Programme: First Interim Report DFE-RR070 London: Department for Education.

Majnemer, A., Shevell, M., Law, M., Poulin, C., & Rosenbaum, P. (2008). Reliability in the ratings of quality of life between parents and their children of school age with cerebral palsy. Quality of Life Research, 17, 1163-1171.

Markham, C. & Dean, T. (2006). Parents' and professionals' perceptions of Quality of Life in children with speech and language difficulty. International Journal of Language & Communication Disorders, 41, 189-212.

Markham, C., van, L. D., Gibbard, D., & Dean, T. (2009). Children with speech, language and communication needs: their perceptions of their quality of life. International Journal of Language & Communication Disorders, 44, 748-768.

McDougall, J., Wright, V., Miller, L., & Nichols, M. (2012). Evaluation of two measures of perceived quality of life when used with youth with chronic conditions. Developmental Medicine and Child Neurology [On-line].

Neuner, B., von, M. S., Krumpel, A., Manner, D., Friefeld, S., Nixdorf, S. et al. (2011). Health-related quality of life in children and adolescents with stroke, self-reports, and parent/proxies reports: cross-sectional investigation. Annals of Neurology, 70, 70-78.

Petrou, S. & Kupek, E. (2009). Estimating preference-based health utilities index mark 3 utility scores for childhood conditions in England and Scotland. Medical Decision Making [On-line].

Pring, T., Flood, E., & Dodd, B. (2012). The working practices and clinical experiences of paediatric speech and language therapists: a national UK survey. International Journal of Language and Communication Disorders 47, 696-708.

Ref Type: Journal (Full)

Rowen, D., Brazier, J., Young, T., Gaugris, S., Craig, B. M., King, M. T. et al. (2011). Deriving a preference-based measure for cancer using the EORTC QLQ-C30. Value in Health, 14, 721-731.

Seid, M., Varni, J. W., & Kurtin, P. S. (2000). Measuring quality of care for vulnerable children: challenges and conceptualization of a pediatric outcome measure of quality. Am.J.Med.Qual., 15, 182-188.

Thomas-Stonell, N. L., Oddson, B., Robertson, B., & Rosenbaum, P. L. (2010). Development of the FOCUS (Focus on the Outcomes of Communication Under Six), a communication outcome measure for preschool children. Developmental Medicine & Child Neurology, 52, 47-53.

Tilford, J. M., Payakachat, N., Kovacs, E., Pyne, J. M., Brouwer, W., Nick, T. G. et al. (2012). Preference-based health-related quality-of-life outcomes in children with autism spectrum disorders: a comparison of generic instruments. Pharmacoeconomics., 30, 661-679.

Varni, J. W., Burwinkle, T. M., Seid, M., & Skarr, D. (2003). The PedsQL 4.0 as a pediatric population health measure: feasibility, reliability, and validity. Ambulatory Pediatrics, 3, 329-341.

Varni, J. W., Seid, M., & Kurtin, P. S. (2001). PedsQL 4.0: reliability and validity of the Pediatric Quality of Life Inventory version 4.0 generic core scales in healthy and patient populations. Med.Care, 39, 800-812.

Vogels, T., Verrips, G. H., Verloove-Vanhorick, S. P., Fekkes, M., Kamphuis, R. P., Koopman, H. M. et al. (1998). Measuring health-related quality of life in children: the development of the TACQOL parent form. Qual.Life Res, 7, 457-465.

Ware, J. E. Jr. & Sherbourne, C. D. (1992). The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Medical care, 30, 473-483.

White-Koning, M., Arnaud, C., Dickinson, H. O., Thyen, U., Beckung, E., Fauconnier, J. et al. (2007). Determinants of child-parent agreement in quality-of-life reports: a European study of children with cerebral palsy. Pediatrics, 120, e804-e814.

Whitehead, S. J. & Ali, S. (2010). Health outcomes in economic evaluation: the QALY and utilities. Br.Med.Bull., 96, 5-21.

Wilkin, D., Hallam, L., & Doggett, M. (1992). Measures of need and outcome for primary health care. Oxford: Oxford University Press.

Willems, D. C., Joore, M. A., Nieman, F. H., Severens, J. L., Wouters, E. F., & Hendriks, J. J. (2009). Using EQ-5D in children with asthma, rheumatic disorders, diabetes, and speech/language and/or hearing disorders. International Journal of Technology Assessment in Health Care, 25, 391-399.

World Health Organisation (2010). International Statistical Classification of Diseases and Health Related Problems, 10th Revision.

Yang, Y., Brazier, J., Tsuchiya, A., & Young, T. A. (2011). Estimating a preference-based index for a 5-dimensional health state classification for asthma derived from the Asthma Quality of Life Questionnaire. Medical Decision Making, 31, 281-291.

Young, A., Gomersall, T., & Bowen, A. (2013). Trial participants' experiences of early enhanced speech and language therapy after stroke compared with employed visitor support: a qualitative study nested within a randomized controlled trial. Clinical Rehabilitation, 27, 174-182.

Table 1. Study characteristics

Author, year	Objective	Study design and setting	Participant characteristics	Inclusion / exclusion criteria
Arkkila 2009	To evaluate the HRQoL of adolescents with an SLI diagnosis	**Design:** cross-sectional survey comparing adolescents with SLI to a group of typically developing peers **Setting:** Department of phoniatrics, Helsinki University Central Hospital.	**N:** 67 adolescents with receptive SLI (48 responded); 235 population controls **Age (mean yr):** 14.2 $\pm$1.44 **Population:** Adolescents with receptive SLI	• Referred to, and examined at the Department of Phoniatrics, Helsinki University Central Hospital • Between 12 to 16 years of age • Diagnosed with receptive SLI (by multidisciplinary team)
Arkkila 2011	To evaluate HRQoL in a clinical sample of children with SLI	**Design:** cross-sectional survey comparing children with SLI to a group of typically developing peers **Setting:** Department of phoniatrics, Helsinki University Central Hospital.	**N:** 60 children with SLI (51 responders) and a group of 244 typically developing schoolchildren **Age (mean yr):** SLI group: 6.28 $\pm$0.86 Typically developing group: NR (however, states the sample was age-matched) **Population:** School-aged children (8 to 11 y/o) with SLI.	• Referred to the Department of Phoniatrics, Helsinki University Central Hospital • Between 8 to 11 years of age • Diagnosed with SLI (by multidisciplinary team) Exclusion: • Hearing impairment • Serious neurological deficits • Metabolic or genetic syndrome
Burgess, 2010	To evaluate the feasibility of using the ASHA's QCL for adolescents with high-functioning autism / asperger syndrome (HFA/AS)	Design: Cross-sectional comparison of adolescents with autism and typically developing peers	**N:** 29 (14 HFA/AS; 15 typically developing) **Age (mean yr):** 16.87 $\pm$ 1.93 (HFA/AS); 15.88 $\pm$ 2.06 (control group) **Population:** Children and adolescents with HFA / AS	• Diagnosis of autism, asperger syndrome, or pervasive developmental disorder, not otherwise specified, by a health professional • Score $\geq$7 on Autism Diagnostic Observation Schedule-Generic • Above average language skills (core language score $\geq$ 85 assessed by CELF • IQ $\geq$ 85 assessed by Kaufman Brief Intelligence Test (2nd ed) • Native English speakers Exclusion • History of neurological disorders
Damiano, 2007	To evaluate the factors that affect the HRQoL of preadolescent children with	**Design:** Cross-sectional telephone / postal survey **Setting:** Unclear clinical	**N:** 104 **Age (mean yr):** 6.5 $\pm$3.1	Inclusion • Diagnosis of oral cleft within first year of life Exclusion • Evidence of additional noncleft structural birth

	nonsyndromic clefts using the Pediatric Quality of Life Inventory instrument	setting; data collected in US	**Population:** Children with nonsyndromic oral clefts	defects, recognised aetiology, or evidence of mental retardation
Flapper 2011	To examine impact of SLI on HRQoL	**Design:** Cross-sectional study comparing children with SLI and typically developing peers **Setting:** Special education schools (n=4); mainstream ambulatory care schools (n=7) in the Northern Netherlands	**N:** 159 (124 children with SLI; 35 matched controls) **Age (mean yr):** NR (range: 5 to 8 years) **Population:** Children and adolescents with SLI	NR – assume any children with SLI
Flapper 2013	To investigate co-morbidity of developmental coordination disorder (DCD) in children with SLI and in investigate impact of DCD on QoL	**Design:** Cross-sectional survey **Setting:** Four special education schools in the Northern Netherlands	**N:** 669 (137 children with SLI; reference group of 532 children from national database) **Age (mean yr):** 6.7 **Population:** Children with SLI	Inclusion • Meets formal criteria for SLI according to DSM criteria (test scores for auditory processing, grammar, or lexical-semantic skills >1.25 SD below the norm) • Age 5 to 8 years Exclusion • IQ score <85 on SON nonverbal intelligence test • Known impairments of speech, hearing, or neurological dysfunction
Keskin 2012	To examine differences in QoL and functional status as reported by adolescents with CP and their mothers	**Design:** Cross-sectional study **Setting:** Department of physiotherapy and rehabilitation, Hacettepe University, Turkey	**N:** 27 mother-adolescent pairs **Age (mean yr):** NR **Population:** Children and adolescents with cerebral palsy	Inclusion • 8 to 18 years of age • Cerebral palsy • Able to complete the QoL questionnaires
Kuhlthau 2010	To examine the relationship between HRQoL and ASD symptoms and behavioural characteristics, and to compare with chronically ill and typically developing peers	**Design:** Cross-sectional study **Setting:** 15 autism centres across the US and Canada	**N:** 286 **Age (mean yr): NR** **Population:** Children and adolescents with ASD	• Confirmed diagnosis of an ASD • Age between 2 years 0 months to 17 years 9 months • Score on ADOS-G must meet or exceed cutoffs for ASD • Fluent English speakers
Lee 2008	To examine QoL and parental concerns in	**Design:** Cross-sectional study	**N:** 6802 (parents / carers of children with autism: 483; parents/carers of children with	• 3 to 17 years of age • Parents responded positively to the question: “has a

	children with autism during early childhood, childhood, and adolescence, compared to children with ADHD	**Setting:** Data from the National Survey of Children's Health (NSCH), conducted in the US in 2003	ADHD: 6319) **Age (mean yr):** NR (children aged 3 to 17 years) **Population:** Children / adolescents with autism / ADHD	doctor / health professional ever told you that [CHILD] has autism." • Comorbid diagnoses, including ADHD, were permitted in this group • Parents who responded positively to the question: "has a doctor / health professional ever told you that [CHILD] has autism." • Comorbid autism was not permitted in this group.
Limbers 2009	To examine feasibility, reliability, and validity of PedsQL in school-aged children with Asperger's syndrome	**Design:** Cross-sectional study **Setting:** Various specialist services for children with ASD	**N:** 22 **Age (mean yr):** 9.25 ±2.15 **Population:** Children and adolescents with Asperger's sndrome	• Confirmed Asperger's syndrome • 6 to 12 years of age
Loy 2010	To examine HRQoL scores from profoundly deaf children and to compare their responses with a matched cohort and their parents	**Design:** Cross-sectional study **Setting:** Summer camp in Estes Park, Colorado and Dallas, Texas	**N:** 88 families (84 followed up) **Age (mean yr):** 10.9 **Population:** Cochlear implant users with profound hearing loss	Inclusion • Documented severe-profound hearing loss • Use of at least one cochlear implant
Majnemer 2008	To compare perspectives of children with cerebral palsy and their parents on the child's QoL	**Design:** Cross-sectional study **Setting:** Secondary care (neurology).	**N:** 48 child-parent pairs **Age (mean yr):** 9.9 ±1.9 **Population:** Children with cerebral palsy	• Children aged 6-12 years • Independent neurological examination to confirm cerebral palsy • Cognitive and language ability to self-report.
McDougall 2012	To evaluate the performance of two QoL measures used with youth and children with various chronic conditions	**Design:** Prospective cohort study **Setting:** Eight children's rehabilitation centres in Ontario, Canada.	**N:** 400 adolescent-caregiver pairs **Age (mean yr):** 14 **Population:** Adolescents with various chronic conditions (incl. CP, communication disorders, autism, TBI)	NR
Mirasola 2006	To examine the self-reported voice-related quality of life of patients diagnosed with paradoxical vocal fold dysfunction (PVFD) using the Pediatric Voice Outcomes Survey (PVOS)	**Design:** Retrospective case series **Setting:** Clinical database in Wisconsin, USA	**N:** 20 **Age (mean yr):** 13.5 **Population:** Patients with PVFD	Inclusion • Inspiratory airway obstruction as primary complaint • Clinical examination: flexible laryngoscopy with videostroboscopy • Other clinical signs including negative metacholine challenge test, dynamic inspiratory obstruction or pulmonary function testing
Petrou &	To estimate preference	**Design:** Nationwide	**N:** total 2236; of which relevant	Children with any disability who had taken part in the survey

Kupek 2009	based HUI multiattribute utility scores associated with a range of childhood conditions	cross-sectional survey **Setting:** UK, "Disability Survey 2000: Survey of Young People With a Disability and Sport"	populations: ASD (n=105); Down's syndrome (n=155); cerebral palsy (n=178); speech disorders (n=25) **Age (mean yr):** ASD: 11.0; Down's syndrome: 12.2; cerebral palsy: 11.6; speech disorders: 10.4 **Diagnosis:** ASD, Down's syndrome, cerebral palsy, speech disorders.	
Thomas-Stonell 2009	To develop an outcome measure that captures real-world changes in preschool children's communication	**Design:** Scale development (3-phase) **Setting:** Four organisations that provide SLT to preschool children in Ontario, Nova Scotia, and Newfoundland and Labrador, Canada.	**N:** 165 **Age (mean yr):** 3.8 $\pm$ 0.91 **Population:** Children with any communication impairment. The most common medical diagnoses were autism (n=15), cerebral palsy (n=2) and Down syndrome (n=2)	NR – assume any children who had attended, and been treated at, the participating centres
Tilford et al., 2012	To describe HRQoL outcomes in children with ASDs; to compare the sensitivity of two generic preference-based instruments for ASD-related conditions	**Design** Cross-sectional prospective collection of outcome data **Setting:** Two sites of the Autism Treatment Network, and an outpatient psychiatric clinic at Columbia University Medical Centre, New York	**N:** 150 child-caregiver pairs **Age (mean yr):** 8.6 $\pm$3.3 **Diagnosis:** Autism spectrum disorders	• Met DSM-IV criteria for a diagnosis of ASD • Age between 4 and 17 years • Family caregiver spoke English
White-Konig 2006	To examine differences in QoL ratings between children with CP and their parents	**Design:** Cross-sectional study **Setting:** 9 regions in 7 European countries (Denmark, France, Germany, Ireland, Italy, Sweden, UK)	**N:** 500 child-caregiver pairs **Age (mean yr):** NR (range: 7 to 13) **Population:** Children with CP	Inclusion • Children with confirmed CP according to criteria of Cans 2000 • Date of birth between 31/7/1991 and 01/04/1997
Willems 2009	To explore variables of the EQ-5D child version in children with chronic conditions (incl. Speech, language and	**Design:** Cross-sectional study **Setting:** Various sites in the Netherlands.	**N:** 194 child-caregiver pairs **Age (mean yr):** For age 7 to 12 group: 10 $\pm$ 1.5 (n=99); for age 12 to 18 group: 15 $\pm$ 1.8 (n=62)	Inclusion • Age 7 to 18 years • Affected with one of the following disorders: asthma, rheumatic disorders, diabetes,

	hearing disorders)		**Population:** Children with various chronic illnesses	speech/language/hearing problems

Abbreviations: ADHD, attention deficit hyperactivity disorder; ASHA, American Speech-Language-Hearing Association; CELF, Clinical Evaluation of Language Fundamentals; CP, cerebral palsy; HRQoL, health-related quality of life; HUI, health utility index; NR, not reported; PODCI, Pediatric Outcome Data Collection Instrument; PVFD, paradoxical vocal fold dysfunction QCL, Quality of Communication Life Scale; QoL, Quality of life; RCT, randomised controlled trial; SLI, specific language impairment; TBI, traumatic brain injury

Table 2. Details of quality of life measure(s) employed in the literature

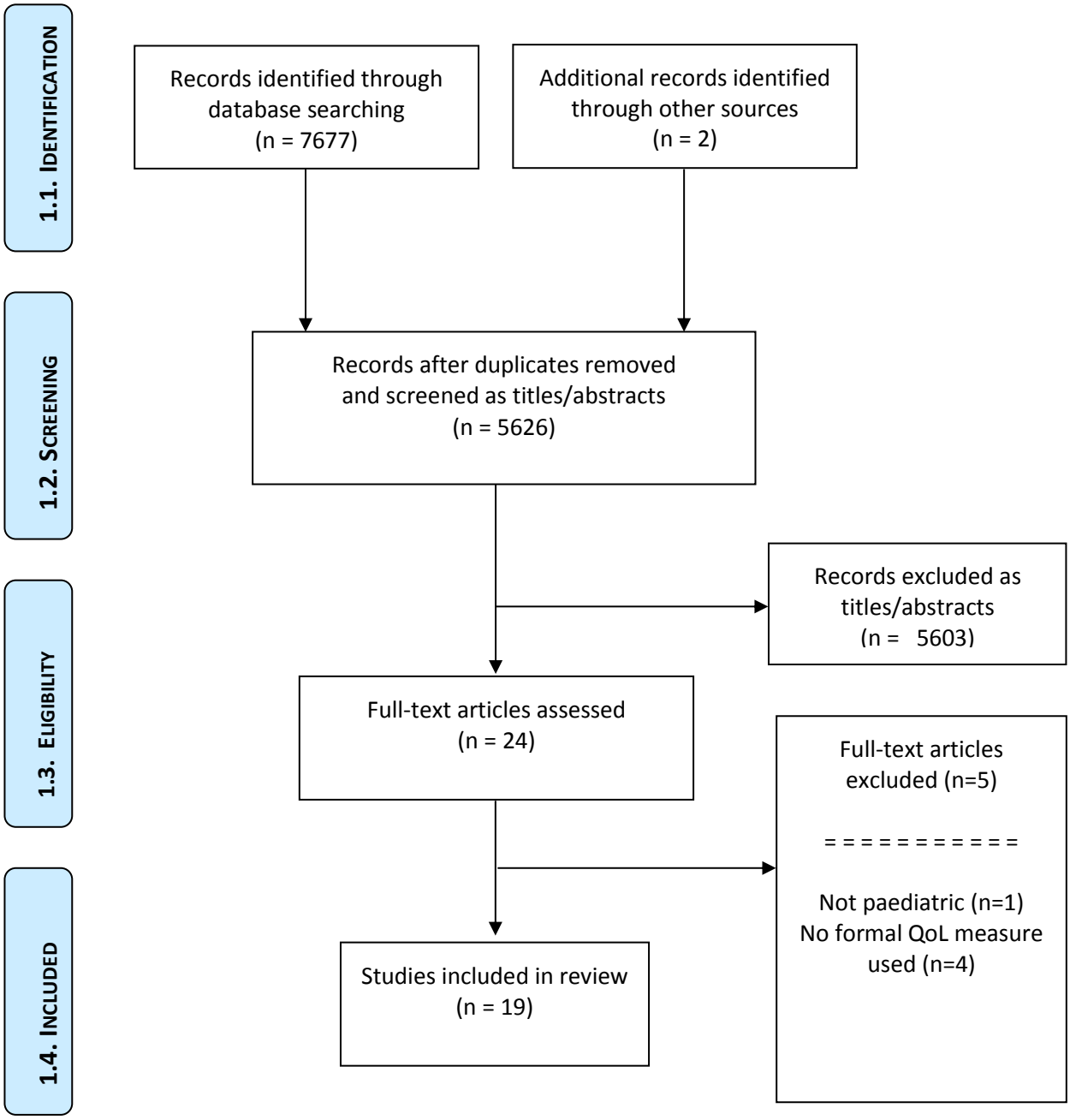
Name of QoL tool used	Studies using the tool	QoL dimensions measured	Scoring of QoL tool
16D	Arkkila 2009	Vitality; breathing; vision; distress; hearing; sleeping; eating; discomfort / symptoms; speech; appearance; school & hobbies; mobility; friends and relations; mental function; elimination; depression	Each domain has five levels. A single index score on a 0-1 scale, representing the overall HRQoL **(0 = being dead, 0.0162 = being unconscious or comatose, 1 = no problems on any dimension = 'full' HRQOL)** is calculated from the health state descriptive system by using a set of population-based preference or utility weights. Such a weight for each level of each dimension is obtained by multiplying the level value by the importance weight of the dimension at that level.
17D	Arkkila 2011	Mobility; Vision; Hearing; Breathing; Sleeping; Eating;	The single index score is on a 0-1 scale, representing overall

		Speech; Excretion; School and hobbies; Learning and memory; Discomfort and symptoms; Depression; Distress; Vitality; Appearance; Friends; Concentration	HRQoL (0 = being dead, 1 = perfect HRQOL) is calculated from the health state descriptive system by using a set of population-based preference or utility weights, which have been elicited from a sample of parents of 8-11 years old school children.
ASHA-QCL (modified for adolescents)	Burgess 2010	Does not attempt to separate different domains on the scale. Includes 18 statements	VAS scale on each statement, with 'yes' aligned with the top, and 'no' aligned with the bottom. Numerical scores are assigned on the basis of the position of the line.
PedsQL	Damiano 2007, Keskin 2012, Majnemer 2008	Physical; Emotional; Social; School	Items are reversed scored and linearly transformed to a 0-100 scale, with higher scores indicating better HRQOL For the **Psychosocial Health Summary Score**, the sum of the items is divided by the number of items answered in the Emotional, Social, and School Functioning Scales. The **Physical Health Summary Score** is the same as the Physical Functioning Scale Score. The **Total Scale Score** can be computed as the sum of all the items over the number of items answered on all the Scales.
TACQOL	Flapper 2011, 2013, Willems 2009	Physical functioning (pain in head/stomach, dizzy, tired, and sleepy); Motor functioning (running, climbing steps, playing, balance); Autonomic functioning (dressing, eating, sports, and hobbies); Cognitive functioning (comprehending, expressing, learning in school); Social functioning (peers / parents); Positive moods (happy and content); Negative moods (sad, angry, occupied)	Each item scored from 0 (worst possible) to 4 (best possible) The scale scores are calculated by a simple summation of the (combined) item scores and a simple correction for missing answers. Combined-item scores are of an ordinal level of measurements only, however, they may be treated as interval scales for the purposes of analysis.
NSCH	Lee 2008	Caring burden; Family outing; Family meals; Religious service attendance; Quit a job; Days missing school; Activity participation; Repeated a grade; Independence; Community service	Each question was scored individually and treated as a separate outcome variable
KINDL-R	Loy 2010	Physical well-being; Emotional well-being; Self-esteem; Family; Friends; School	Five-point likert scale for each item (never, seldom, sometimes, often, all the time) Subscales are summed for an overall score and transformed into a 100-point scale, with 0 representing worst possible QoL and 100 representing best possible QoL
BMLSS;	McDougall 2012	Single-factor structure with four dimensions: • intrinsic (myself, overall life), • social (friendships, family life), • external (work, where I live), • perspective (financial situation, future Prospects)	Each item scored on a 7-point likert scale (0: terrible to 6: delighted). BMLSS sum score is presented as a %/100

SLSS	McDougall 2012	Unidimensional factor structure: global life satisfaction in children .	Each item scored on a 4-point likert scale (never = 1; to almost always = 4). Negatively-keyed items must be reverse scored. Hence, higher scores indicate higher levels of life satisfaction throughout the scale. Domain and total scores are made comparable by summing the item responses and dividing by the number of domain (or total) items.
PVOS	Mirasola 2006	NR	5-point scale on three of the four items, and 3-point scale on one item. Raw scores are converted to a scale of 0 (worst) to 100 (best)
HUI3	Petrou & Kupek 2009; Tilford 2012	Vision; hearing; speech; ambulation; dexterity; emotion; cognition; pain/discomfort (each having five or six functional levels)	The classification of eight domains into five or six levels can be used to define 972,000 unique health states. HUI is scored using single- and multi-attribute utility functions. A utility score is derived between 0.00 (dead) to 1.00 (best possible QoL). Negative scores are permitted to recognise states worse than death, with the worst possible score being -0.36
FOCUS	Thomas-Stonell 2009	Phase 1 & 2 of testing (factor analyses) indicated a single underlying construct: "real-world communication outcomes"	
Quality of Well-Being Self-Administered (QWB-SA)	Tilford 2012	Chronic symptoms; acute symptoms and mental health; mobility; physical activity; social activity	A health utility score between 0 (for dead) to 1.0 (for perfect health) can be generated using population-level preference weights. The scoring algorithm may be obtained on request from the scale authors (Seiber et al. 2008)
Kidscreen	White-Konig 2006	• Physical Well-being, • Psychological Well-being, • Moods and Emotions, • Self-Perception, • Autonomy, • Parent Relations and Home Life • Social Support and Peers • School Environment, • Social Acceptance (Bullying), • Financial Resources.	Item responses are summed for each domain and a score out of 100 is computed, with higher scores indicating better QoL.
EQ-5D	Willems 2009	• Mobility • Self-care • Usual activities • Pain/discomfort • Anxiety/depression	**EQ-5D:** Five levels per dimension, from 1 (no problems to level 5 (extreme problems). 3125 possible health states can be derived by combining one level from each of the five dimensions. These can be converted into a single index score ranging from 0.00 (dead) to 1.00 (best possible QoL), by weighting the scores according to population-level preferences.

Abbreviations: ASHA, American Speech-Language-Hearing Association ; BMLSS, Brief Multidimensional Life Satisfaction Scale ; CI, cochlear implant ; Focus, Focus on the Outcomes of Communication Under Six ; HUI, Health Utilities Index ; NSCH, National Survey of Children's Health ; PedsQL, Pediatric Quality of Life Inventory ; PODCI, Pediatric Outcomes Data Collection Instrument ; PVOS, Pediatric Voice Outcomes Survey ; QCL, Quality of Communication Life Scale ; SLSS, Students' Life Satisfaction Scale ; TACQOL, TNO-AZL Questionnaire for Children's Health-Related Quality of Life ; VAS, visual analogue scale

Figure 1: PRISMA flow chart



Appendix 1. Web of Science search strategy

# 26	1,377	#24 AND #21 Refined by: Languages=(ENGLISH) *Databases=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH Timespan=All Years*
# 25	1,456	#24 AND #21 *Databases=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH Timespan=All Years*
# 24	376,823	#23 OR #22 *Databases=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH Timespan=All Years*
# 23	186,624	Topic=((outcome* NEAR/3 (indicator* or assess* or measur* or scale* or score* or index or indices or status or questionnaire* or instrument* or monitor*))) *Databases=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH Timespan=All Years*
# 22	210,255	Topic=(quality of life) *Databases=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH Timespan=All Years*
# 21	27,423	#20 AND #10 *Databases=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH Timespan=All Years*
# 20	81,990	#19 OR #18 OR #17 OR #16 OR #15 OR #14 OR #13 OR #12 OR #11 *Databases=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH Timespan=All Years*
# 19	9,577	Topic=(communication disorder*) *Databases=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH Timespan=All Years*
# 18	8,393	Topic=(language deficit*) *Databases=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH Timespan=All Years*
# 17	9,668	Topic=(speech disorder*) *Databases=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH Timespan=All Years*
# 16	17,152	Topic=(learning disorder*) *Databases=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH Timespan=All Years*
# 15	5,437	Topic=(language delay*) *Databases=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH Timespan=All Years*
# 14	14,015	Topic=(language disorder* or language development disorder*) *Databases=SCI-EXPANDED, SSCI, CPCI-S, CPCI-*

SSH Timespan=All Years

# 13	28,267	Topic=(learning difficult*) *Databases=SCI-EXPANDED, SSCI, CPCI-S, CPCI-SSH Timespan=All Years*
# 12	5,169	Topic=(communication impairment*) *Databases=SCI-EXPANDED, SSCI, CPCI-S, CPCI-SSH Timespan=All Years*
# 11	10,324	Topic=(language impairment*) *Databases=SCI-EXPANDED, SSCI, CPCI-S, CPCI-SSH Timespan=All Years*
# 10	1,820,124	#9 OR #8 OR #7 OR #6 OR #5 OR #4 OR #3 OR #2 OR #1 *Databases=SCI-EXPANDED, SSCI, CPCI-S, CPCI-SSH Timespan=All Years*
# 9	11,798	Topic=(schoolchild*) *Databases=SCI-EXPANDED, SSCI, CPCI-S, CPCI-SSH Timespan=All Years*
# 8	22,486	Topic=(teen*) *Databases=SCI-EXPANDED, SSCI, CPCI-S, CPCI-SSH Timespan=All Years*
# 7	33,868	Topic=(pre-school* or preschool*) *Databases=SCI-EXPANDED, SSCI, CPCI-S, CPCI-SSH Timespan=All Years*
# 6	148,883	Topic=(boy* or girl*) *Databases=SCI-EXPANDED, SSCI, CPCI-S, CPCI-SSH Timespan=All Years*
# 5	40,776	Topic=(baby or babies) *Databases=SCI-EXPANDED, SSCI, CPCI-S, CPCI-SSH Timespan=All Years*
# 4	580,779	Topic=(young or youth*) *Databases=SCI-EXPANDED, SSCI, CPCI-S, CPCI-SSH Timespan=All Years*
# 3	292,903	Topic=(infant*) *Databases=SCI-EXPANDED, SSCI, CPCI-S, CPCI-SSH Timespan=All Years*
# 2	31,495	Topic=(p?ediatri*) *Databases=SCI-EXPANDED, SSCI, CPCI-S, CPCI-SSH Timespan=All Years*
# 1	1,075,791	Topic=(child*) *Databases=SCI-EXPANDED, SSCI, CPCI-S, CPCI-SSH Timespan=All Years*

Appendix 2. Data extraction form

DATA EXTRACTION	
REVIEW DETAILS	
Author, year	
Objective	
INTERVENTIONS (IF RELEVANT)	
Description	
T1: Intervention group	
T2: Control group	
Length of follow up	
STUDY CHARACTERISTICS	
Type of study (RCT, scale validation, etc)	
Numbers included in the study	
Site (health care setting, country)	
POPULATION CHARACTERISTICS	
Target population (describe)	
Inclusion / exclusion criteria (n)	
Recruitment procedures used (participation rates if available)	
CHARACTERISTICS OF PARTICIPANTS	
Age (mean yr.)	

Female (n, %)	
Previous speech therapy?	
Other information	
QoL Measurement	
Tool used	
Scale dimensions (list)	
Individual scale items (list)	
Self or proxy measurement?	
Preference weighting used Y/N? If Y, describe.	
Validation?	
Scoring method (describe)	