Preferences in the Design and Delivery of Neurodevelopmental Follow-Up Care for Children: A Systematic Review of Discrete Choice Experiments

Pakhi Sharma, Sanjeewa Kularatna, Bridget Abell, Karen Eagleson, Linh K Vo, Ureni Halahakone, Sameera Senanayake, Steven M McPhail

1Australian Centre for Health Services Innovation and Centre for Healthcare Transformation, School of Public Health and Social Work, Queensland University of Technology, Brisbane, QLD, Australia; 2Queensland Paediatric Cardiac Service, Queensland Children’s Hospital, Brisbane, QLD, Australia; 3Faculty of Medicine, The University of Queensland, Brisbane, QLD, Australia; 4Digital Health and Informatics Directorate, Metro South Health, Brisbane, QLD, Australia

Correspondence: Pakhi Sharma, Australian Centre for Health Services Innovation and Centre for Healthcare Transformation, School of Public Health and Social Work, Queensland University of Technology, 60 Musk Avenue, Kelvin Grove, Brisbane, QLD, 4059, Australia, Tel +61 401829045, Email p33.sharma@hdr.qut.edu.au

Abstract: Neurodevelopmental disorders are a significant cause of morbidity. Early detection of neurodevelopmental delay is essential for timely diagnosis and intervention, and it is therefore important to understand the preferences of parents and clinicians for engaging with neurodevelopmental surveillance and follow-up care. Discrete choice experiment (DCE) may be an appropriate method for quantifying these preferences. This review systematically examined how DCEs have been designed and delivered in studies examining neurodevelopmental care of children and identified the preferred attributes that have been reported. PubMed, Embase, CINAHL, and Scopus databases were systematically searched. Studies were included if they used DCE to elicit preferences for a neurodevelopmental follow-up program for children. Two independent reviewers conducted the title and abstract and full-text screening. Risk of bias was assessed using a DCE-specific checklist. Findings were presented using a narrative synthesis. A total of 6618 records were identified and 16 papers were included. Orthogonal (n=5) and efficient (n=5) experimental designs were common. There was inconsistent reporting of design-related features. Analysis was primarily completed using mixed logit (n=6) or multinomial logit (n=3) models. Several key attributes for neurodevelopmental follow-up care were identified including social, behavioral and emotional support, therapy, waiting time, and out-of-pocket costs. DCE has been successfully used as a preference elicitation method for neurodevelopmental-related care. There is scope for improvement in the design and analysis of DCE in this field. Nonetheless, attributes identified in these studies are likely to be important considerations in the design and implementation of programs for neurodevelopmental care.

Keywords: preferences, discrete choice experiment, attributes, neurodevelopment, follow-up

Introduction
Neurodevelopmental disorders encompass a range of conditions that affect early development and functioning of the brain. Difficulties may arise in social, cognitive, or emotional domains and include diagnoses such as attention-deficit /hyperactive disorder (ADHD) and autism spectrum disorder (ASD). Consequently, neurodevelopmental disorders can cause significant morbidity in children with subsequent impact on families and incur substantial costs to healthcare and education systems. A survey from the United States reported that 15% of children aged 3 to 17 years were affected by neurodevelopmental disorders. Along with genetic and environmental factors, critical illnesses early in life, including congenital heart disease and prematurity, are also a known risk factor for neurodevelopmental delays. Early detection of neurodevelopmental delay is important for timely diagnosis and intervention, and it is therefore important to...
understand the preferences of parents and clinicians for engaging with neurodevelopmental services and follow-up care.\textsuperscript{8–10} The use of the term “neurodevelopmental follow-up care” in this manuscript refers to the care provided to children and their families over time after diagnosing the neurodevelopmental disorder or delay in order to monitor the medical and developmental outcomes and therefore provide suitable support.\textsuperscript{11,12} Knowing what parents and clinicians prefer, in terms of neurodevelopmental follow-up care, is likely to aid in the design of services to enhance uptake and acceptability.

There are different methods of understanding patient preferences, both qualitative and quantitative. Some studies have used surveys, interviews, and focus groups to identify preferences in neurodevelopmental follow-up care;\textsuperscript{13–17} however, the experimental design and analysis methods in these studies were not robust for quantifying preferences in this context. Regarding care for children with neurodevelopmental disorders, there may be many important attributes and the weighted preferences for each of these cannot be readily captured and disentangled using these methods for the purpose of informing health service design. Consequently, quantitative choice methods may be suitable to use in this context as they may be more reliable at predicting authentic behavior\textsuperscript{18,19} and can provide a robust method to capture and quantify the most relevant preferences.

Discrete choice experiment (DCE) is a quantitative choice method, where individuals are presented several hypothetical health scenarios (choice sets), each containing several alternatives with different attributes for the individuals to choose from.\textsuperscript{20} A DCE can be designed using the following key steps: identifying whether an experiment is labelled or unlabelled; identifying and finalising attributes; identifying the attribute levels; identifying the number of choice sets in experimental design; selecting a suitable experimental design strategy; conducting a pilot study; conducting the main study; and analysing the results using a suitable analysis method.\textsuperscript{21,22} DCE is considered a robust technique that uses the ‘framework of rational choice’, which infers that people are likely to select the option that will yield maximum benefit or utility when given a set of choices. In this context, for an individual i, making a choice j, utility (U) is described as a function of characteristics of the choices (Zj) and the characteristics of the people making the choice (Xi) and an error term denoting unobserved attributes of choices and individuals (eij).\textsuperscript{23}

$$U_{ij} = F(X_i, Z_j, e_{ij})$$

In this equation, F is assumed to be a linear function (hence this is also known as linear random utility model) and can be represented as:

$$U_{ij} = \beta Z_j + \gamma Z_j X_{ij} + e_{ij}$$

where β and γ are the utilities (also known as parameter estimates) associated with the choice features and interaction of choice features and individual features, respectively.

DCE, being a stated preference technique, is useful in determining weighted preferences of various stakeholders, including patients, parents and clinicians in a healthcare decision-making situation, which is not possible via other common methods such as interviews, focus groups, and surveys.\textsuperscript{24} Additionally, they estimate monetary and non-monetary values, for example, willingness to pay (WTP), willingness to accept (WTA), and probability scores,\textsuperscript{21,25} which can be beneficial to include in a neurodevelopmental follow-up care as it may help in better understanding the approximate expenditure for future follow-up care.

Although the application of DCEs in fields of healthcare is relatively new, the use of DCEs in studies of neurodevelopmental follow-up care has recently increased. However, methodological limitations in experimental design and analysis methods, as well as sub-optimal reporting in publications have been observed in this field.\textsuperscript{26–29} In order to advance the quality and strength of conclusions that can be drawn from DCEs in this field, it is important to understand the methodological features that have been used, as well as identify potential neurodevelopmental care attributes appropriate for use in future preference studies. This systematic review aimed to examine how DCEs have been designed and delivered in studies for neurodevelopmental follow-up care of children, and to identify preferred attributes of these models of care for the purpose of informing neurodevelopmental follow-up care health service design.
Methods
This systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses report (PRISMA) guidelines. A protocol for this review was registered with PROSPERO (Record ID = CRD42022325685).

Data Sources, Search Strategy, and Screening
On 08 February 2022, we searched PubMed, Embase, CINAHL, and Scopus databases. The search strategy was developed using existing literature and expert assistance from researchers experienced with systematic reviews in this field. A combination of Medical Subject Heading (MeSH) terms and keywords were used for groups of terms related to “children”, “patient preference” and “neurodevelopmental follow-up care”. Search syntax and count of returned titles for each string for all databases are provided in Online Resource 1 (Tables S1–S6). An advanced Google search was also conducted with the first 20 pages screened; however, no additional relevant papers were identified.

All records obtained through the searches were uploaded to EndNote version 20 and deduplicated prior to screening. These records were uploaded to Rayyan, a web-based systematic review management platform, for screening. Two authors independently completed the title and abstract screening of all the records identified from the databases (PS and UH). Any disagreements were resolved by consulting a third author (SK). After the title and abstract screening, full-text publications were independently screened by another pair of reviewers (PS and SS) and disagreements were resolved by consulting SK.

Study Selection
Original research studies were included if (1) the population was children, infants or adolescents, (2) a DCE methodology was used, and (3) the context was neurodevelopmental follow-up care or programs for neurodevelopmental disorders (for example, ADHD, ASD, dyslexia, and cerebral palsy). Studies using any other preference methods including surveys and interviews were excluded. Publications were excluded if they were reviews, meta-analysis, case studies, case reports, conference abstracts, letters to editor, guidelines, and commentaries. Non-human studies were also excluded. No limits were applied on publication date or language.

Data Extraction
Two authors were involved in data extraction for included studies, one extracted data (PS) and the other checked the extracted data (SS) for all studies. Any unresolved disagreements were resolved by consulting a third author (SK). Missing data were reported for the final analysis and interpretation. All information was extracted into a Microsoft Excel Workbook.

The following data were extracted: publication details (author name, year, country, title, and journal name); study details (aims/objectives, study design, dates of recruitment, primary and secondary outcome); neurodevelopmental disorder and children’s details (type of neurodevelopmental disorder/delay, inclusion criteria for children in the study, age, and sample size); respondent details (respondent type, age, sample size, response rate, and recruitment strategy); DCE methodological details (for example, attributes included for a model of care, experimental design, analysis method, and software used), results, conclusions, limitations, and the advantages and disadvantages of the studies.

Risk of Bias Assessment
We assessed risk of bias in the included studies with a widely used checklist initially created by Lancsar and Louviere in 2008, and modified by Mandeville in 2014. Domains include choice task design, experimental design, conduct, and analysis. Two reviewers (PS and LV) independently applied the checklist to each study and recorded judgements for each item. Any unresolved disagreements were resolved by consulting a third author (SK).

Data Analysis
The data were summarised descriptively and tabulated for each DCE study included in the review. A narrative synthesis was completed for demographic characteristics, attribute selection process, experimental design, analysis method, and the identified attributes.
Results
A total of 6618 papers were identified; after removing duplicates, 6361 remained. During screening, most studies were excluded as they were not performed in the context of neurodevelopment follow-up or were not focussed on children. From the 49 publications that progressed to full-text screening, 16 studies were included in the review.26–29,35–46 The screening and selection process is reported in PRISMA diagram in Figure 1 Preferred Reporting Items for Systematic Reviews and Meta-Analyses report.

Characteristics of Included Studies
Table 1 summarises the important study characteristics of the included DCE publications. Studies were conducted from the years 2008 to 2022 mostly in the United States of America (n=5)28,29,36,41,42 and Canada (n=3)26,27,40 or both (n=2),44,45 followed by European countries (n=5),37–39,43,46 and Australia (n=1).35 ADHD was the most frequently studied neurodevelopmental disorder (n=7),37–39,42,44–46 followed by children’s mental health (CMH) (n=2)26,27 and ASD (n=1).29 Parents (n=9)27,37–40,42,44–46 or caregivers (n=4)28,35,36,41 were the main DCE respondent type. However, two studies also reported professionals (including child and youth worker, social workers, psychologist, psychiatrist, early

Identification of studies via databases

![Diagram showing the steps of selecting studies: Identification, Screening, and Included]

Figure 1 Preferred Reporting Items for Systematic Reviews and Meta-Analyses report.
childhood education providers, nurses) and one study reported teachers as respondents. The number of participants in studies ranged from 38 to 1194 with the median sample size being 241.

Methodological Characteristics of Included DCE Studies

Table 2 describes the attribute selection, choice set generation, and experimental designs used in each included DCE study.

Attribute Selection Process

Authors used a variety of methods to select attributes including focus group discussions (n=8) and literature reviews (n=6) and sometimes a combination of these and other methods. Studies included two (n=8), three (n=7), or five (n=1) alternatives. The number of attributes included in studies ranged from 3 to 20 and the maximum levels included were 5. Attributes can either be quantitative (numerical) or qualitative (categorical). Usually, categorical variables need to be coded to use in utility functions. Most studies in this review lacked reporting on the method used to code categorical variables (n=11). Of those that did report this, two used dummy coding and two used effects coding. Only two studies reported utility specifications, which stated using alternative specific constant.

Table 1 Demographic Characteristics of the Included DCE Studies (n=16)

<table>
<thead>
<tr>
<th>Countries</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>United States of America</td>
<td>5</td>
</tr>
<tr>
<td>Europe</td>
<td>5</td>
</tr>
<tr>
<td>Canada</td>
<td>3</td>
</tr>
<tr>
<td>Canada, United States of America</td>
<td>2</td>
</tr>
<tr>
<td>Australia</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Neurodevelopmental Disorder/Delay</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attention-deficit/hyperactivity disorder (ADHD)</td>
<td>7</td>
</tr>
<tr>
<td>Children's mental health (CMH)</td>
<td>2</td>
</tr>
<tr>
<td>Fragile X syndrome (FXS)</td>
<td>1</td>
</tr>
<tr>
<td>Autism Spectrum Disorder (ASD)</td>
<td>1</td>
</tr>
<tr>
<td>Idiopathic developmental disability (DD)</td>
<td>1</td>
</tr>
<tr>
<td>Intellectual disability (ID)</td>
<td>1</td>
</tr>
<tr>
<td>Prader-Willi syndrome (PWS)</td>
<td>1</td>
</tr>
<tr>
<td>Coexisting cognitive/intellectual and an emotional/behavior/developmental disability</td>
<td>1</td>
</tr>
<tr>
<td>Communication disabilities</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Respondent Type</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parents</td>
<td>9</td>
</tr>
<tr>
<td>Caregivers</td>
<td>4</td>
</tr>
<tr>
<td>Professionals</td>
<td>2</td>
</tr>
<tr>
<td>Teachers</td>
<td>1</td>
</tr>
</tbody>
</table>

Abbreviation: N, count of studies.
<table>
<thead>
<tr>
<th>Author and Year</th>
<th>Attributes Selection Process</th>
<th>Number of Attributes Included</th>
<th>Maximum Number of Levels within Each Attribute</th>
<th>Number of Choice Sets per Participant</th>
<th>Blocks</th>
<th>Experimental Design Type</th>
<th>Type of Fractional Factorial</th>
<th>Algorithm (If Efficient Design)</th>
<th>Draws (If Bayesian Design)</th>
<th>Choice Task Format (Labelled/ Unlabelled)</th>
<th>Experimental Design Approach</th>
<th>Categorical Variables Coding Method</th>
<th>DCE Pilot Test (Yes/ No)</th>
<th>Priors Used (Yes/ No)</th>
<th>Method of Prior Selection</th>
<th>Survey Administration Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arora 2019&lt;sup&gt;23&lt;/sup&gt;</td>
<td>Pilot study and similar experiments identified in literature</td>
<td>5</td>
<td>5</td>
<td>16</td>
<td>8</td>
<td>Fractional factorial</td>
<td>Bayesian D-efficient design</td>
<td>Modified Federov algorithm</td>
<td>500 Halton draws</td>
<td>Labelled</td>
<td>Forced choice approach</td>
<td>Not stated</td>
<td>Yes</td>
<td>Yes</td>
<td>Existing literature&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Online</td>
</tr>
<tr>
<td>Cross 2016&lt;sup&gt;24&lt;/sup&gt;</td>
<td>Literature review and expert opinion</td>
<td>6</td>
<td>3</td>
<td>9</td>
<td>6</td>
<td>Fractional factorial</td>
<td>Bayesian D-efficient design</td>
<td>D-optimal algorithm</td>
<td>NA</td>
<td>Unlabelled</td>
<td>Forced choice approach</td>
<td>Effects</td>
<td>No</td>
<td>No</td>
<td>Not stated</td>
<td>NA</td>
</tr>
<tr>
<td>Cunningham 2008&lt;sup&gt;25&lt;/sup&gt;</td>
<td>FGD</td>
<td>20</td>
<td>4</td>
<td>30</td>
<td>No</td>
<td>Fractional factorial</td>
<td>Not stated</td>
<td>NA</td>
<td>NA</td>
<td>Unlabelled</td>
<td>Forced choice approach</td>
<td>Not stated</td>
<td>Yes</td>
<td>Not stated</td>
<td>NA</td>
<td>Email and paper-based</td>
</tr>
<tr>
<td>Cunningham 2009&lt;sup&gt;26&lt;/sup&gt;</td>
<td>FGD</td>
<td>14</td>
<td>4</td>
<td>20</td>
<td>No</td>
<td>Fractional factorial</td>
<td>Not stated</td>
<td>NA</td>
<td>NA</td>
<td>Unlabelled</td>
<td>Forced choice approach</td>
<td>Not stated</td>
<td>No</td>
<td>No</td>
<td>Not stated</td>
<td>NA</td>
</tr>
<tr>
<td>dosReis 2020&lt;sup&gt;27&lt;/sup&gt;</td>
<td>FGD, debriefing session</td>
<td>6</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>Fractional factorial</td>
<td>Orthogonal design (type: balanced)</td>
<td>NA</td>
<td>NA</td>
<td>Unlabelled</td>
<td>Forced choice approach</td>
<td>Effects</td>
<td>No</td>
<td>Not stated</td>
<td>NA</td>
<td>Email, telephone</td>
</tr>
<tr>
<td>Fegert 2011&lt;sup&gt;28&lt;/sup&gt;</td>
<td>Literature research, qualitative survey, and FGD</td>
<td>6</td>
<td>2</td>
<td>Not stated</td>
<td>No</td>
<td>Fractional factorial</td>
<td>Orthogonal design</td>
<td>NA</td>
<td>NA</td>
<td>Unlabelled</td>
<td>Forced choice approach</td>
<td>Not stated</td>
<td>No</td>
<td>No</td>
<td>Not stated</td>
<td>NA</td>
</tr>
<tr>
<td>Glenngård 2013&lt;sup&gt;29&lt;/sup&gt;</td>
<td>Previous study from the UK</td>
<td>5</td>
<td>4</td>
<td>8</td>
<td>2</td>
<td>Fractional factorial</td>
<td>Orthogonal design</td>
<td>NA</td>
<td>NA</td>
<td>Unlabelled</td>
<td>Forced choice approach</td>
<td>Not stated</td>
<td>Yes</td>
<td>Not stated</td>
<td>NA</td>
<td>Paper</td>
</tr>
<tr>
<td>Hugh 2022&lt;sup&gt;30&lt;/sup&gt;</td>
<td>Questionnaire</td>
<td>5</td>
<td>Not stated</td>
<td>5</td>
<td>No</td>
<td>Not stated</td>
<td>Not stated</td>
<td>NA</td>
<td>NA</td>
<td>Labelled</td>
<td>Forced choice approach</td>
<td>Not stated</td>
<td>No</td>
<td>Yes</td>
<td>Not stated</td>
<td>NA</td>
</tr>
<tr>
<td>Mühlbacher 2009&lt;sup&gt;31&lt;/sup&gt;</td>
<td>Literature research and FGD</td>
<td>6</td>
<td>2</td>
<td>Not stated</td>
<td>No</td>
<td>Fractional factorial</td>
<td>Orthogonal design</td>
<td>NA</td>
<td>NA</td>
<td>Unlabelled</td>
<td>Not stated</td>
<td>Not stated</td>
<td>No</td>
<td>Yes</td>
<td>Not stated</td>
<td>NA</td>
</tr>
<tr>
<td>Nafiss 2014&lt;sup&gt;32&lt;/sup&gt;</td>
<td>Literature review and expert opinion</td>
<td>7</td>
<td>3</td>
<td>18</td>
<td>No</td>
<td>Fractional factorial</td>
<td>Orthogonal design (type: array)</td>
<td>NA</td>
<td>NA</td>
<td>Unlabelled</td>
<td>Forced choice approach</td>
<td>Not stated</td>
<td>Yes</td>
<td>Not stated</td>
<td>NA</td>
<td>Online</td>
</tr>
<tr>
<td>Raiger 2009&lt;sup&gt;33&lt;/sup&gt;</td>
<td>Literature review and expert opinion</td>
<td>3</td>
<td>4</td>
<td>16</td>
<td>No</td>
<td>Fractional factorial</td>
<td>Bayesian D-efficient design</td>
<td>Hierarchical Bayesian algorithm</td>
<td>2000 (type of draws not stated)</td>
<td>Unlabelled</td>
<td>Opt-out approach</td>
<td>Dummy</td>
<td>Yes</td>
<td>No</td>
<td>Not stated</td>
<td>NA</td>
</tr>
</tbody>
</table>

<sup>a</sup> Existing literature refers to the use of literature reviews or similar studies to identify attributes and design parameters.
<table>
<thead>
<tr>
<th>Study</th>
<th>Title</th>
<th>Patient Advisory Board</th>
<th>N</th>
<th>T</th>
<th>F</th>
<th>Fractional factorial</th>
<th>D-efficient design</th>
<th>Number of attributes</th>
<th>Type</th>
<th>Priors</th>
<th>Value of priors</th>
<th>Cash compensation attribute</th>
<th>Pilot study</th>
<th>Other Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tsai 2021</td>
<td>Communications among a community advisory board including patients, caregivers, and clinicians</td>
<td>4</td>
<td>2</td>
<td>12</td>
<td>No</td>
<td>Fractional factorial</td>
<td>Not stated</td>
<td>Not stated</td>
<td>Unlabelled</td>
<td>Forced choice approach</td>
<td>Reverse</td>
<td>Yes</td>
<td>Yes</td>
<td>Not stated</td>
</tr>
<tr>
<td>Waschbusch 2017</td>
<td>Survey and expert discussions</td>
<td>18</td>
<td>3</td>
<td>27</td>
<td>No</td>
<td>Fractional factorial</td>
<td>Not stated</td>
<td>Not stated</td>
<td>Unlabelled</td>
<td>Forced choice approach</td>
<td>Dummy</td>
<td>No</td>
<td>No</td>
<td>NA</td>
</tr>
<tr>
<td>Webb 2019</td>
<td>BWS, FGD, interview</td>
<td>9</td>
<td>4</td>
<td>12</td>
<td>5</td>
<td>Fractional factorial</td>
<td>Not stated</td>
<td>NA</td>
<td>Labelled</td>
<td>Forced choice approach</td>
<td>Not stated</td>
<td>Yes</td>
<td>No</td>
<td>NA</td>
</tr>
<tr>
<td>Wymbs 2016</td>
<td>FGD</td>
<td>20</td>
<td>4</td>
<td>30</td>
<td>No</td>
<td>Fractional factorial</td>
<td>Not stated</td>
<td>Not stated</td>
<td>Unlabelled</td>
<td>Forced choice approach</td>
<td>Not stated</td>
<td>No</td>
<td>No</td>
<td>NA</td>
</tr>
<tr>
<td>Wymbs 2017</td>
<td>FGD</td>
<td>20</td>
<td>4</td>
<td>30</td>
<td>No</td>
<td>Fractional factorial</td>
<td>Not stated</td>
<td>Not stated</td>
<td>Unlabelled</td>
<td>Forced choice approach</td>
<td>Not stated</td>
<td>No</td>
<td>No</td>
<td>NA</td>
</tr>
</tbody>
</table>

**Notes:** a. Value of priors: each of the care types = 0.1, cash compensation attribute = −0.1. b. Pilot study.

**Abbreviations:** DCE, discrete choice experiment; FGD, focus group discussion; BWS, best-worst scale; NA, not applicable.
Generating Choice Sets
In this review, there were more unlabelled (n=13) than labelled choice sets (n=3). Most studies used a forced choice approach (n=14), where respondents have no choice but to answer that choice set. Only one study used an opt-out approach,\(^{38}\) and one study did not report this information.\(^{38}\) The total number of choice sets in included studies ranged from 8 to 609 but the number of choice sets presented per participant ranged from 3 to 30. None of the studies explicitly stated whether they used a homogenous (respondents are shown the same choice sets) or heterogeneous (respondents are shown only a subset of choice sets by dividing into blocks or groups) design matrix. However, most studies did not block their designs (n=11); therefore, we assumed that they used a homogenous design matrix.

Experimental Design Strategy
In this review, all studies used fractional factorial design except one that did not state the type of experimental design.\(^{29}\) Orthogonal design (n=5) was a common design: one study performed an orthogonal array\(^{39}\) and one an orthogonal balance.\(^{36}\) Efficient designs were also identified (n=5), of which three were D-efficient designs and two were Bayesian D-efficient. Of the Bayesian D-efficient designs, one study used a Modified Federov algorithm and 500 Halton draws\(^{35}\) and the other used a Hierarchical Bayesian algorithm with 2000 draws (type of draws was not stated).\(^{40}\) Six studies did not report which fractional factorial category they used.

Study Type and DCE Administration
In this review, half of the studies conducted a pilot study, and the other half did not. One of the studies included in the review was itself a pilot study.\(^{36}\) Only two studies used priors in this review,\(^{35,41}\) while the others did not state using or not using them. Arora et al obtained prior values of +0.1 and −0.1 via existing literature\(^{35}\) and Tsai et al used zero valued priors.\(^{41}\) Arora et al also reported that their sample was obtained from a broader clinical trial while the remaining studies did not mention any information on sample size. Most DCE surveys were administered online or via email (n=9), one was paper based, few were both (n=5), and one was via both telephone and email.

Analysis Methods in Included DCE Studies
The analysis methods used in each DCE are presented in Table 3. Various models were used in the DCE studies (Figure 2 Types of analysis methods identified in the included studies [n = 16]). Most of the studies used a single model (n=13) and the rest used multiple models. Of those that used one model, mixed logit was common (n=6), followed by multinomial logit (MNL) (n=3), and random-effects logit (n=2). Studies with multiple models used mixed logit and latent class analysis (LCA) (n=2) and MNL and LCA (n=1). Two of the mixed logit model studies extended their model specification to generalised MNL (GMNL)\(^{35}\) and heteroskedastic MNL,\(^{41}\) which makes different assumptions about parameter distributions. For model estimation, only five studies have stated using Hierarchical Bayes estimation method.\(^{26,27,42,44,45}\) While it can be assumed that others have used maximum likelihood estimation, this was not reported explicitly. Of those studies that used LCA, classes were selected using log-likelihood, Akaike information criterion (AIC) or Bayesian information criterion (BIC). Preference heterogeneity was reported in 10 studies. Studies accounted for explained, unexplained, and scale heterogeneity. Only three studies performed monetary evaluations and reported willingness to pay (WTP) (n=2)\(^{37,40}\) and willingness to accept (WTA) (n=1).\(^{35}\)

Outcomes of Included DCE Studies: Preferred Attributes
In DCE studies, the attributes preferred by parents or caregivers and clinicians differed from each other. For parents and caregivers, the social, behavioral, and emotional situation of the child were important. Effectiveness and side effects of medication, receiving a genetic diagnosis, lesser waiting time, and spending less income were also identified as important attributes by parents. Individual and group parent training was highlighted as important by some parents. One study indicated that parents preferred treatments with behavior therapy and the desire to avoid medication.\(^{42}\) For clinicians, a child’s controlling behavior, ability to self-care, decision-making, and future communication and expression skills were important attributes. A study also found that professionals prefer active learning materials with parenting groups and therapist coaching.
Table 3 Characteristics of the Analysis Methods and Most Preferred Attributes Identified in the Included DCE Studies (n=16)

<table>
<thead>
<tr>
<th>Author and Year</th>
<th>Analysis Methods (Model Type)</th>
<th>Preference Heterogeneity (Yes/No)</th>
<th>Utility Specification</th>
<th>Variable Distribution</th>
<th>Post Estimation Calculations</th>
<th>Most Preferred Attributes Identified (Results)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arora 2019\textsuperscript{35}</td>
<td>Mixed logit and LCA</td>
<td>Yes</td>
<td>ASC</td>
<td>Normal distribution</td>
<td>WTA</td>
<td>Social support. WTA = $A35.96</td>
</tr>
<tr>
<td>Cross 2016\textsuperscript{38}</td>
<td>Mixed logit</td>
<td>Yes</td>
<td>Not stated</td>
<td>Not stated</td>
<td>No</td>
<td>Controlling behavior, caring for oneself</td>
</tr>
<tr>
<td>Cunningham 2008\textsuperscript{27}</td>
<td>MNL and LCA</td>
<td>Yes</td>
<td>Not stated</td>
<td>Not stated</td>
<td>No</td>
<td>Step-by-step solutions to behavioral or emotional problems, weekly meetings with other parents and coaching calls from a therapist</td>
</tr>
<tr>
<td>Cunningham 2009\textsuperscript{26}</td>
<td>LCA</td>
<td>Yes</td>
<td>Not stated</td>
<td>Not stated</td>
<td>No</td>
<td>Active learning materials with parenting groups and therapist coaching calls supporting the knowledge transfer process</td>
</tr>
<tr>
<td>dosReis 2020\textsuperscript{36.a}</td>
<td>MNL</td>
<td>Not stated</td>
<td>Not stated</td>
<td>Not stated</td>
<td>No</td>
<td>Using fewer medications, maintaining decision-making authority</td>
</tr>
<tr>
<td>Fegert 2011\textsuperscript{46}</td>
<td>Random-effects logit</td>
<td>Not stated</td>
<td>Not stated</td>
<td>Not stated</td>
<td>No</td>
<td>Social situation, emotional state</td>
</tr>
<tr>
<td>Glenngård 2013\textsuperscript{37}</td>
<td>Random-effects logit</td>
<td>Not stated</td>
<td>Not stated</td>
<td>Not stated</td>
<td>WTP</td>
<td>Effectiveness, side-effects, and number of dosing per day. WTP = €790 for adolescents. Could not calculate for children.</td>
</tr>
<tr>
<td>Hugh 2022\textsuperscript{29}</td>
<td>MNL</td>
<td>Yes</td>
<td>Not stated</td>
<td>Not stated</td>
<td>No</td>
<td>Naturalistic intervention (NI)</td>
</tr>
<tr>
<td>Mühlbacher 2009\textsuperscript{38}</td>
<td>Logit/Probit</td>
<td>Not stated</td>
<td>Not stated</td>
<td>Not stated</td>
<td>No</td>
<td>Social situation, emotional state</td>
</tr>
<tr>
<td>Nafees 2014\textsuperscript{39}</td>
<td>MNL</td>
<td>Yes</td>
<td>Not stated</td>
<td>Not stated</td>
<td>No</td>
<td>Degree of symptom control</td>
</tr>
<tr>
<td>Reiger 2009\textsuperscript{40}</td>
<td>Mixed logit</td>
<td>Yes</td>
<td>Not stated</td>
<td>Normal distribution: attribute representing the number of children tested. Fixed: other attributes</td>
<td>WTP</td>
<td>Receiving a genetic diagnosis, waiting less time and spending less income. WTP = CDN$1118</td>
</tr>
</tbody>
</table>

(Continued)
Table 3 (Continued).

<table>
<thead>
<tr>
<th>Author and Year</th>
<th>Analysis Methods (Model Type)</th>
<th>Preference Heterogeneity (Yes/No)</th>
<th>Utility Specification</th>
<th>Variable Distribution</th>
<th>Post Estimation Calculations</th>
<th>Most Preferred Attributes Identified (Results)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tsai 2021</td>
<td>Mixed logit</td>
<td>Yes</td>
<td>Not stated</td>
<td>Not stated</td>
<td>No</td>
<td>Without hyperphagia = accept greater weight gain and a higher risk of skin rash. With hyperphagia = accept a higher risk of liver damage</td>
</tr>
<tr>
<td>Waschbusch 2011</td>
<td>Mixed logit and LCA</td>
<td>Yes</td>
<td>Not stated</td>
<td>Not stated</td>
<td>No</td>
<td>Treatments with behavior therapy, desire to avoid medication</td>
</tr>
<tr>
<td>Webb 2019</td>
<td>Mixed logit</td>
<td>Yes</td>
<td>ASC</td>
<td>Normal distribution</td>
<td>No</td>
<td>Future skills and abilities, child’s determination and persistence, communication ability with AAC, and receptive and expressive language</td>
</tr>
<tr>
<td>Wymbs 2016</td>
<td>Mixed logit</td>
<td>Not stated</td>
<td>Not stated</td>
<td>Not stated</td>
<td>No</td>
<td>Individual parent training</td>
</tr>
<tr>
<td>Wymbs 2017</td>
<td>Mixed logit</td>
<td>Not stated</td>
<td>Not stated</td>
<td>Not stated</td>
<td>No</td>
<td>Group parent training (when individual PT is not available)</td>
</tr>
</tbody>
</table>

**Note:** *Pilot study.*

**Abbreviations:** LCA, latent class analysis; MNL, multinomial logit; ASC, Alternative specific constant; WTA, Willingness to accept; WTP, Willingness to pay.
Table 3 highlights only the most important attributes (as a result of model outcome) in DCE studies. The range of attributes initially included by authors in all of the studies is presented in Online Resource 2 (Table S7).

Risk of Bias and Quality of Included Studies
An assessment of the reporting quality and risk of bias of included studies is provided in Table 4. There were high risks in the choice task design specially in the inclusion of an opt-out or status quo option or justification of forced choice. Not justifying why a forced choice was included in the choice set may obligate the participants to choose something they are not interested in. Moderate to high risks were also identified in the experimental design of the study which may hinder with the neurodevelopment follow-up care outcome as the choice tasks may be inadequately designed. The conduct of discrete choice experiments was typically of low risk with an exception in the ‘response rate sufficient to minimise response bias’ section. Studies either had a low response rate or did not report it at all. A high response rate may ensure better generalisability of outcomes. The analysis of discrete choice experiments was also typically of low risk. Econometric models selected were particularly appropriate for choice task design.

Discussion
The primary aim of this systematic review was to synthesize the methods and findings of DCE studies conducted in the context of neurodevelopmental follow-up for children. Our review identified that most studies had been conducted over the past decade. This aligns with a literature review in 2019 reporting a 20% increase in DCE use over the preceding 10 years. 

Most included studies focussed on ADHD and ASD which are the two most common neurodevelopmental disorders observed in childhood. Overall, the findings of this review suggested variability in the methods used by the studies and highlighted some important attributes from the perspective of parents and carers. Most importantly, the design of a DCE is crucial to the conduct and quality of the study, and it includes several steps that require attention to detail. Here, we discuss the key shortcomings of the DCE studies included in this review and suggest opportunities for improvement.

Most studies identified in this review were unlabelled. Even though the research question may dictate whether an experiment is labelled or unlabelled, evidence suggests that a participant’s choice may sometimes be influenced by labels when they are present. The utility function (described as the measurement of consumers’ preferences) for an unlabelled experiment is identical or generic, but for a labelled experiment the utility function should be specified. Two out of three labelled studies in this review used label-specific constants. The importance of using utility functions accurately has been highlighted in previous research as it is crucial to identify the relevant importance of an attribute over other, and in turn, better understand respondents’ behavioral responses and satisfaction with the overall model.
Table 4 Quality Assessment of the Included DCE Studies (n=16)

<table>
<thead>
<tr>
<th></th>
<th>Choice Task Design</th>
<th>Experimental Design</th>
<th>Conduct</th>
<th>Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Choice of Attributes and Levels Grounded in Qualitative Work with Target Population</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Araja 2019&lt;sup&gt;13&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cross 2016&lt;sup&gt;24&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cunningham 2008&lt;sup&gt;8&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cunningham 2009&lt;sup&gt;29&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dosReis 2020&lt;sup&gt;36&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fegert 2011&lt;sup&gt;14&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glenngård 2013&lt;sup&gt;37&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hugh 2022&lt;sup&gt;43&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mühlbacher 2009&lt;sup&gt;38&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nafees 2014&lt;sup&gt;40&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reiger 2009&lt;sup&gt;41&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tsai 2021&lt;sup&gt;42&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Waschbusch 2011&lt;sup&gt;43&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Webb 2019&lt;sup&gt;44&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wymbs 2016&lt;sup&gt;45&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wymbs 2017&lt;sup&gt;46&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: White, criteria met (manuscript's text sufficiently confirmed the criteria); Black, criteria not met (not reported or no evidence to justify the criteria in text); Grey, not enough evidence to justify the criteria.
This review reported four studies conducting a partial profile, which means they showed only a subset of attributes in each choice task due to large number of attributes. Evidence suggests that there is significant impact of the number of alternatives, attributes and levels shown to the respondents in a DCE on their behavioral responses. A smaller number of attributes reduce cognitive burden on respondents, although quantity may not be as important as the relevance of the attribute. Consequently, partial profiling may be beneficial in studies comprising a large number of attributes and is an important methodological consideration for future studies in this field.

Most studies did not report blocking their experimental design and are likely to have used a homogenous design. Prior evidence indicates that heterogeneous designs are advantageous because they provide more information and allow blocking that may reduce cognitive burden for respondents. However, a homogeneous design may be appropriate if few parameters are to be estimated. Seven studies in this review used a large number of choice sets. Although the number of choice sets is typically determined by the total number of parameters to be estimated in the choice model, there has been inconsistency in prior recommendations regarding the number needed. Hensher et al suggested using 4 to 16 choice tasks. This is in contrast to studies that have demonstrated the number of choice sets may have minimal impact on findings. Nonetheless, regardless of the methodological nuance required to address specific research questions, potential fatigue caused to the respondents remains a key consideration when designing studies in this field.

Design strategy was typically not clearly reported in the included studies. This limited our ability to draw strong conclusions regarding the most appropriate approach for conducting neurodevelopmental follow-up DCEs. On the one hand, efficient designs may be better in that they capture maximum information, calculate reliable parameter estimates with smaller sample sizes, and can be used if any attribute levels are unusual (unrealistic or impossible) and dominant. On the other hand, orthogonal designs may be better at encompassing the attribute space, and are less technically demanding to conduct. However, in the studies included in this review, capturing maximum and reliable information seemed to be prioritised in this context of investigating preferences for neurodevelopmental follow-up care.

Half of the studies identified in the review conducted pilot studies and only two studies reported information on priors. For the remaining studies, this is a notable limitation as there is no indication of accuracy regarding how the parameter values were obtained. Previous studies suggest that conducting a pilot study is important for generating feedback about the DCE and to obtain parameter priors so that a better design and impression of likely parameter estimates can be obtained for the main study. Moreover, if informative priors are estimated, this will assist with minimum required sample size calculations.

The analysis methods used by DCEs in this review each have strengths and limitations. Mixed logit model was a common analysis method among studies in this review, followed by MNL. MNL (also known as conditional logit) has been used and modified by many researchers since its introduction in 1973. However, due to advancement in software packages and increasing demand, more variations and complexities were added in estimating models which likely contributed to mixed logit and latent class analysis models receiving more attention in the literature. Furthermore, MNL accounts for explained preference heterogeneity; and mixed logit, random-effects logit, and latent class models account for unexplained preference heterogeneity. Preference heterogeneity refers to different respondents having different preferences in the same choice scenario which a recent study suggests is important in health-related DCEs. Preference heterogeneity was reported in 10 studies in this review, two of which also reported scale heterogeneity (observed in scale parameters). This indicates that the authors acknowledged different behaviors may influence neurodevelopmental follow-up care.

**Implications for Future Research**

The nature of the research question as well as the quality and quantity of data are influential and interdependent considerations when determining the most appropriate experimental DCE design. This review has highlighted opportunity for methodological improvements at all stages of the design and analysis process to improve robustness of DCEs in this field. During the initial design stage of a DCE, careful consideration of the attribute selection process, generation of choice sets, and strategising a suitable experimental design among future studies in the field is likely to aid in optimising the accuracy and informativeness of findings. Similarly, appropriate prior selection and sample size...
calculation are key methodological features that should be purposefully devised. It is also important for future studies to consider response behaviors that may be related to unexplained heterogeneity that can be accounted for through appropriate analyses including mixed logit or latent class models. More monetary value calculations could help in building better health service models and offer stakeholders an understanding about an estimate expenditure for follow-up care.

Strengths and Limitations of the Review

Although this review considered a broad scope of literature with no date or language limitations, which may be considered a strength, the included publications required an appropriate level of methodological description to achieve the aim of this review. This meant that some publication types, including conference abstracts, case studies, and letters to editor, were considered unsuitable and excluded from the review and consequently findings from this review are not inclusive of information reported in other publication forms. While this was appropriate for addressing the intended study aims, it is noteworthy that among the included journal manuscripts unclear reporting still often contributed to uncertainty in our understanding of some important components of DCE methodology used. Due to the limited number of studies and heterogenous sample characteristics across diagnostic sub-groups, it was also not possible for separate analyses to be completed across sub-groups of neurodevelopmental disorders. Rather, the scope of the present review was limited to providing an overview of preferences for informing a healthcare service design perspective, that is not necessarily neurodevelopment diagnosis specific, regarding overall approaches to neurodevelopment-related follow-up care. Due to the nature of neurodevelopmental disorders and individual circumstances, differences in preferences may exist both among consumers within the same diagnosis category and between diagnostic groups. However, health system design internationally has not typically involved the establishment of different health systems and services for individual diagnostic groups, but rather services that provide neurodevelopmental follow-up are typically designed to provide services for a range of at-risk children with or without specific diagnoses at the time of entering the service. While we have tried to be inclusive of a range of diagnosis and disorders in the scope of searches for this study, the available literature was not inclusive of all diagnostic groups that may access services for neurodevelopmental follow-up care. Consequently, while findings addressed the study aims regarding the design and delivery of DCEs examining neurodevelopmental care of children and identified the preferred attributes that have been reported, the findings arising from studies included in this review should not be interpreted as representing the preferences of all consumers and families who may access neurodevelopmental follow-up care.

Conclusion

In summary, several aspects of DCEs were assessed in neurodevelopmental follow-up care. DCE is an appropriate preference method as it allows the estimation of weighted preferences, uncertainties, preference heterogeneity, the relevance of an attribute over another, and offers significant preference information beneficial to patient-centred care. The attributes identified from studies in this review may increase awareness of important components of future follow-up programs. Furthermore, when discussing management options, findings from this review may contribute to a broader understanding of potential areas for focussed discussions that may be important from the perspective of children who require neurodevelopmental support, along with the needs of their families.

Abbreviations

DCE, Discrete Choice Experiment; ADHD, Attention-deficit/hyperactive disorder; ASD, Autism Spectrum Disorder; WTP, Willingness to Pay; WTA, Willingness to Accept; PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses report; MeSH, Medical Subject Headings; MNL, Multinomial Logit; LCA, Latent Class Analysis; GMNL, Generalised Multinomial Logit; AIC, Akaike Information Criterion; BIC, Bayesian Information Criterion.
Key Points

Key attributes for a follow-up care for children needing neurodevelopmental support, along with the needs of their families were identified. A range of opportunities exist for improving robustness of experimental design of discrete choice experiments in this field.

Acknowledgement

The authors would like to acknowledge Associate Professor Sanjeewa Kularatna and Professor Steven M McPhail for their invaluable guidance and support as the senior authors of this paper.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This work is part of a program of research funded through a Medical Research Future Fund (MRFF) Congenital Heart Disease Grant (ARGCHDG0035) 2020-2024. This funder did not have any role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Disclosure

The authors report no conflicts of interest in this work.

References

2. Jeste SS. Neurodevelopmental behavioral and cognitive disorders. CONTINUUM. 2015;21(3):690–714. doi:10.1212/01.CON.0000466661.89908.3c
63. Collins AT, Bliemer MC, Rose JM. Constrained stated choice experimental designs. 2014.