
Article:
Webb, EJD orcid.org/0000-0001-7918-839X, Meads, D orcid.org/0000-0003-1369-2483, Eskyte, I et al. (8 more authors) (2018) A Systematic Review of Discrete-Choice Experiments and Conjoint Analysis Studies in People with Multiple Sclerosis. Patient, 11 (4). pp. 391-402. ISSN 1178-1653

https://doi.org/10.1007/s40271-017-0296-y

A systematic review of discrete choice experiments and conjoint analysis studies in people with Multiple Sclerosis

Edward J.D. Webb[1,2] David Meads[2] Ieva Eskyte[3] Natalie King[2] Naila Dracup[2] Jeremy Chataway[4] Helen Ford[5] Joachim Marti[6] Sue H. Pavitt[7] Klaus Schmierer[8,9] Ana Manzano[10]

Abstract

Background: Multiple sclerosis (MS) is a chronic disabling, inflammatory and degenerative disease of the central nervous system which, in most cases, requires long term disease modifying treatment (DMT). The drugs used vary in efficacy and adverse effect profile. A number of studies have used attribute-based stated preference methods, primarily to investigate patient preferences for initiating or escalating DMT.

Aims: To conduct a systematic review of attribute-based stated preference studies in people with MS to identify common methods employed, and to assess study quality with reference to the specific challenges of this disease area.

[1] Corresponding author, email: e.j.d.webb@leeds.ac.uk, tel: +44 (0)113 343 7356
[2] Leeds Institute for Health Science, University of Leeds
[3] School of Dentistry, University of Leeds
[4] Queen Square Multiple Sclerosis Centre, Department of Neuroinflammation, UCL Institute of Neurology, University College London , London, WC1B 5EH, UK
[5] Leeds Teaching Hospitals NHS Trust
[6] Centre for Health Policy, Imperial College London
[7] Dental Translational & Clinical Research Unit, School of Dentistry, University of Leeds
[8] Blizard Institute (Neuroscience), Barts and The London School of Medicine & Dentistry, Queen Mary University of London, London, UK
[9] Clinical Board Medicine (Neuroscience), The Royal London Hospital, Barts Health NHS Trust, London, UK
[10] School of Sociology and Social Policy, University of Leeds

Methods: We conducted a systematic search for studies related to attribute-based stated preference and MS in multiple databases including Cochrane and MEDLINE. Studies were included if they were published in a peer-reviewed journal, were on the topic of MS and used a survey methodology that measured stated preferences for attributes of a whole. Analysis was conducted using narrative synthesis and summary statistics. Study quality was judged against the ISPOR conjoint analysis checklist.

Results: We identified 16 relevant articles reporting 17 separate studies, all but one focussing on DMTs. The majority of studies were discrete choice experiments. The study quality was generally high, but recommendations are made to improve: 1) sample size considerations, 2) documentation of justification of survey design choices, 3) incorporation of qualitative approaches for attribute and level selection to better involve patients, and 4) better reporting of experimental practice. The effects of DMTs on reproduction and the impact of presenting risk and uncertainty were identified as neglected research topics. The ISPOR conjoint analysis checklist was not found to be suitable to assess study quality.

Conclusion: Attribute-based stated preference is a useful method to examine preferences of people with MS in their choice of DMT. However, there is a need for further research embracing the methodological recommendations identified, particularly greater use of qualitative methods in attribute development.

Key points:

- We conducted a systematic review of discrete choice experiments, conjoint analysis and other attribute-based stated preference studies in Multiple Sclerosis.
- Areas for improvement in future studies are: sample size considerations, documentation of justification for design choices, greater use of qualitative methods for attribute/level development and better reporting of experimental procedures.
- Effect of treatment on reproduction and the influence of risk perception were identified as understudied topics.

Acknowledgements

This work was funded by the UK Multiple Sclerosis Society Grant no. 30. JC acknowledges the UK National Institute for Health Research (NIHR) University College London Hospitals Biomedical Research Centres Funding scheme and University College London (UCL).
SP acknowledges Dental Translational and Clinical Research Unit (DenTCRU), part of Leeds NIHR Clinical Research Facility (CRF).

Contributions:
EW, DM, IE, SP, AM conceived the study.
NK and ND designed and ran searches.
EW and DM selected papers for review.
EW, DM, IE and AM developed the data extraction form.
EW, DM, IE and AM analysed the papers.
EW wrote the manuscript.
All authors revised the manuscript and approved it for publication.

Disclosures: EW, DM, IE, SP, JM and AM have received support from the UK Multiple Sclerosis society. JC has received support from the Efficacy and Mechanism Evaluation Programme and Health Technology Assessment Programme (NIHR); and the UK Multiple Sclerosis Society. In the last three years he has been a local principal investigator for trials in multiple sclerosis funded by: Receptos, Novartis and Biogen Idec, and has received an investigator grant from Novartis outside this work. He has taken part in Advisory Boards/consultancy for Roche, Merck, Biogen, MedDay and Apitope. KS has been a PI of trials sponsored by MedDay, Novartis, Roche, Teva; has been involved in trials sponsored by BIAL, Biogen, Canbex, Cytokinetics, Genzyme; received speaking honoraria from, and/or served in an advisory role for Biogen, Merck, Merck Inc., Novartis, Roche, Teva; has been supported for attendance of meetings by Genzyme, Merck and Novartis, and received research support from Biogen and Novartis.

1. Introduction

Multiple sclerosis (MS) is a chronic neurological disease and the commonest non-traumatic cause of acquired disability in young adults in the Western World [1]. Mean age of onset is 30 years with over two thirds of patients being female [2]. The aetiology of the disease is not fully understood, but it is known to be an inflammatory demyelinating disorder of the central nervous system [3]. Most people with MS (PwMS) experience two clinical phases, initially relapsing-remitting MS (RRMS) followed by a phase with gradual accumulation of disability – secondary progressive MS (SPMS) [4]. Natural history data suggest the clinical phenotype switch from RRMS to SPMS usually occurs about 10-15 years after onset [5, 6]. Whilst the clinical hallmark of RRMS are relapses followed by a variable degree of remission, SPMS is characterised by disability that may affect numerous functions including gait, balance, vision, cognition, and continence [7]. In about 10% of PwMS the disease is progressive from onset – primary progressive MS (PPMS) [8].

Treatments for PwMS fall broadly into two categories: 1) Symptomatic treatments intended to alleviate specific symptoms PwMS experience and 2) Disease modifying treatments (DMTs) intended to alter the natural course of MS, i.e. reducing the frequency and severity of relapses, and slowing of functional deterioration [9]. For DMTs to be effective, PwMS need to commit to long term interventions, in many cases requiring regular administration of tablets, injections or infusions.

Currently, 14 DMTs are available for the treatment of people with relapsing-remitting MS (PwRMS) whilst only one has been approved for people with progressive MS. The drugs used vary in efficacy, adverse event profile, mode of delivery and monitoring burden[10].

The increasing number of DMT options creates uncertainty in treatment selection. There is little information about how PwMS choose DMTs once an MS diagnosis has been established. Several of the most effective DMTs are associated with an increased risk of adverse effects, including life-threatening infections and secondary autoimmunity. Patients have to trade-off these potential negative consequences with perceived benefits (reduced relapse rate and disability accrual, maintained or improved quality of life). Such decisions can be challenging at any time, but may be particularly difficult soon after diagnosis when PwMS are coming to

terms with the presence of a chronic condition, have less knowledge about MS, how it will progress and impact on their quality of life.

The choice of DMT depends greatly on individual preference, requiring weighing-up and trading-off different attributes. For example, a decision has to be made whether a reduction in the probability of relapses outweighs the risk of a serious side-effect. Attribute-based stated preference (AbSP) techniques, such as discrete choice experiments (DCE), best-worst scaling (BWS) and conjoint analysis[1], may elicit such trade-offs between the individual attributes which make up a choice object, and are hence ideal for investigating the DMT preferences of PwMS [11]. Given the number of DMTs is still expanding, another advantage of using AbSP is that it gives an insight into patient attitudes towards potential treatments that are not yet available, and give an indication to those developing and trialling new drugs on what combination of attributes would be acceptable to PwMS.

The above describes why MS is a fertile grounds for AbSP research. However, it is also uniquely challenging. MS is categorized into distinct clinical phenotypes and is a progressive disease with a wide range of symptoms, resulting in a highly individual experience for PwMS. The benefits of treatment are probabilistic: no drug is effective in every case, so every decision to start a DMT represents, to some extent, a gamble. In addition, the clinical endpoints of trials measuring the efficacy of DMTs can be difficult to translate into meaningful terms for PwMS. The trade-offs PwMS have to consider when choosing between different treatments involve all aspects of their current and future lives. For example, they need to consider how much negative impact from side effects on quality of life and daily routine is acceptable in order to potentially slow down accrual of disability several years later. Moreover, many PwMS experience effects on cognition of their disease [12-16], which may impact on their ability to give considered responses in surveys.

While general reviews of DCEs and BWS in health exist [17-20], none has specifically examined MS. Given the significant opportunities and challenges listed above, as well as a

[1] Note that conjoint analysis is sometimes used as an umbrella term to refer to all studies which measure trade-offs between attributes, and is also sometimes treated, including by papers in this review, as synonymous with DCE. Both these uses are incorrect [7]. DCEs (and in some interpretations, BWS [8]) are grounded in random utility theory and are choice based, making them conceptually different from conjoint analysis, which is grounded in mathematical psychology and is ranking/rating based. Therefore, in this paper we use the term attribute-based stated preference (AbSP).

recent rise in the number of relevant studies, a review focussing on this disease area is timely. This paper systematically reviews AbSP studies focusing on experiment design and conduct and suggests recommendations for improvement.

2. Methods and materials

2.1. Search strategy

Comprehensive literature searches were developed by an information specialist using Medline, Embase, PsycINFO, CINAHL, Cochrane Libraries and the Web of Science Core Collections from database inception to 11 July 2017. Search concepts included multiple sclerosis and synonyms[1] and AbSP related terms such as DCE, BWS, max diff and conjoint analysis. Subject headings and free text words were identified for use in the search concepts by the information specialist and project team members. Further terms were identified and tested from known relevant papers. Before running the searches, all search strategies were peer reviewed by a second information specialist using the Peer Review of Electronic Search Strategies (PRESS) checklist [21, 22].[2]

The results of the database searches were stored and de-duplicated in an EndNote library. Further relevant studies were sought by citation searching (forwards and backwards) of the included studies.

The searches identified 328 records. Once duplicates were removed there were 214 records. Citation searches identified 0 records. Two authors (EW and DM) reviewed abstracts and selected 38 for fulltext review. The same two authors then selected for final inclusion articles which were (1) published in a peer-reviewed journal, (2) dealt exclusively or primarily with MS and (3) used an AbSP methodology in any part of the article. An AbSP methodology was defined as any method that uses quantitative data to examine preferences for attributes of a whole. Disagreements were resolved by consensus discussion between EW and DM. This resulted in 16 articles which reported 17 studies.

[1] Specifically disseminated sclerosis and encephalomyelitis disseminata.
[2] Please see Appendix A for full search strategies and date range.

2.2. Data extraction and analysis

Table 1 lists the studies included for final analysis. Data was extracted using the form in appendix 0 by one author (EW). The data extraction form was developed by four authors (EW, DM, IE and AM) with the aim of focusing on study design features (study type, country of origin, participant inclusion criteria, sample size, attribute and levels identification and development) to assess whether studies had used best practice. Given the small numbers of studies it was not thought effective to pilot the form, however minor revisions were made during the process of data extraction. Analysis was performed using a narrative synthesis approach [23]. Detailed consideration was given to attribute development and presentation of information about probabilities, as these are often mentioned as key neglected areas in AbSP practice [24].

The quality of each study was scored by one author (EW) by using the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) conjoint analysis checklist [25]. This contains 30 items which were given a count of 1 if a study reported considering at least some aspect of this item, and 0 if it did not. The final count for a study was then the sum of its counts for each item. A secondary aim of this study was to assess the suitability of using the checklist to assess the quality of AbSP studies for future reviews. This was done by considering whether it showed variation in overall counts and counts for individual items, as well as whether it revealed important issues previously unconsidered.

Narrative synthesis of the identified studies was performed by several authors (EW, AM, IE, DM). Statistical analysis of numerical data was done by computing summary statistics using R version 3.3.1.

3. Results

All but one study examined patient preferences for DMTs, with the remaining study (Rosato, et al. [26]) looking at quality of life for PwMS. This focus is not surprising, as decisions about DMTs are of vital importance to PwMS and feature a mixture of benefits and risks, making AbSP an ideal quantitative tool with which to study the decision making of PwMS.

Fifteen studies (88.2%) were funded by pharmaceutical companies. One study, Rosato et al. [26] received funding from a public health authority and a charitable foundation, and one, Kremer et al. [27] states that the authors received no specific funding.

Figure 2 shows the publication of MS AbSP studies over time. An upward time trend can clearly be seen, with the first appearing in 2009 and nine studies (52.9%) being published since 2016.

3.1. Study type

Table 1 lists the AbSP method used by each study. The majority (nine, 52.9%) were DCEs. Two studies (11.8%) were BWS, of which one used case 1 (Kremer et al. [27]) and one used case 2 (Lynd et al. [28]).

3.2. Survey population

Table 1 lists in which country each study was carried out. The US was the most common country with seven (41.2%), with a further nine (52.9%) spread across Europe (Germany, Italy, the Netherlands, Spain, UK) and a single study in Canada.

3.3. Diagnoses

Seven studies (41.2%) included anyone with a diagnosis of MS, and seven (41.2%) included only patients with a diagnosis of RRMS. Of nine (52.9%) studies that clearly indicated diagnoses to exclude, all excluded PPMS. Seven studies (41.2%) required the diagnosis to be confirmed by a physician, seven (41.2%) relied on self-reported diagnoses, and in three studies (17.6%) it remained unclear how the diagnosis was established.

3.4. Development of attributes and levels

Most studies drew on existing literature in medicine and the social sciences (14, 82.4%) and/or healthcare professionals (12, 70.6%) to develop attributes and levels. Few used qualitative methods to elicit views of PwMS, with only two studies (11.8%) (Lynd, et al. [28], Kremer, et

al. [27]) employing focus groups. Seven studies (41.2%) used interviews at some point in the design stage, typically to refine an existing survey rather than as a basis for attribute development. Two out of these seven (28.6%) did not state how many interviews were carried out (Reed Johnson et al. [29], Poulos et al. [30]), and the average number for the remaining five is 10.3. Two studies (11.8%) (Wicks, et al. [31] studies 1 and 2) did not state how attributes and levels were developed.

3.5. Survey design

Figure 3a shows the number of attributes used by each study. The median number of attributes included in studies is six, in line with the typical number included in AbSP studies in health [18, 19, 32]. The minimum number of attributes was three, the maximum was 27. The median number of levels for each attribute is three, with a maximum of seven and a minimum of two.

Fourteen studies (82.4%) used a fractional factorial design, and two studies (11.8%) (Wicks, et al. [31] studies 1 and 2) do not state the type of design used. Five studies (39.4%), all DCEs, selected their designs based on efficiency, and two (11.8%) explicitly reported using the criterion of D-efficiency.[1] One study (Utz, et al. [34]) used a custom design with a contrast between DMT administration via pill or injection in every choice and all combinations of other attributes presented. Seven studies (41.2%) did not state which criteria were used to construct their design. The statistics program SAS (SAS Institute) was the most popular tool used to construct study designs with four studies (23.5%) and Sawtooth (Sawtooth Software) was the second most popular with three (17.6%). Five studies (29.4%) did not report how their designs were constructed.

Only two out of 16 studies (12.5%) on DMT choice included an opt-out option (Carlin, et al. [26]) or justified why an opt-out was not included (Wilson, et al. [18]).

A concern in designing AbSP surveys is how many choice tasks can be included without the survey becoming a burden to participants [25]. There was considerable variation in survey length, as can be seen in Figure 3b, which illustrates the number of choice tasks per subject in

[1] A D-efficient design is one constructed using an algorithm to maximize the determinant of the information matrix and is commonly used in experimental design construction [33].

each study. The median number was 12, with a standard deviation of 14.1, which is broadly in line with the wider AbSP health literature [18, 19, 32]. However, the number of choices is only one aspect of burden. Another aspect is the complexity of the task. For example, although Utz, et al. [34] presented 64 choice tasks, with only two options and three attributes each, the choices were relatively simple. Several studies increased the total number of choice tasks without increasing the burden on participants by using several different versions of the survey, with the median number being four.

Two studies (11.8%) (Wicks, et al. [31] studies 1 and 2) did not report how many decisions participants made, making it difficult to assess whether the burden was appropriate or not, nor did they report how many survey versions were used. Nine (53.9%) assessed response quality and/or its impact on results (e.g. by presenting the same choice twice, giving a dominated option or eliciting if participants picked the same alternative for every question ("straight-lining") which can be used as an indication both of understanding and that burden was not excessive.

Figure 3 illustrates the sample sizes obtained for final analysis and it can be seen that there is considerable variation. The median sample size was 189 (s.d. 162). Only a single study (Wicks, et al. [31] study 1) reported explicit power calculations and only six (35.3%) reported other sample size considerations such as "rules of thumb".

Eleven studies (64.7%) were administered online and five (29.4%) were administered using pen and paper. Only two studies (Wilson, et al. [35] Bottomley, et al. [36]) reported, in line with item 7.2 of the ISPOR conjoint analysis checklist [25], a justification of the chosen mode of administration.

3.6. Attributes

The attributes used by each study were collated and placed in 13 categories by one author (EW). All attributes were assigned to at least one category and some were assigned to two categories (for example "route and frequency of administration" was classed both as route of administration and frequency of administration).

Among the most common attributes were effect on relapse (13, 76.5%), effect on progression (12, 70.6%), as well as severe side effects (12, 70.6%) and mild side effects (13, 76.5%). Also common were route (10, 58.8%) and frequency (13, 76.5%) of administration. Only four (23.5%) looked at monitoring of treatment, and another four (23.4%) included further miscellaneous aspects of administration. Six studies (35.3%) explored attributes related to the alleviation of MS symptoms. Three (17.6%) included attributes explicitly related to quality of life, one of which (Rosato et al [26]) looked specifically at PwMS' valuation of health-related quality of life. Four (23.5%) included attributes related to MRI scans. Two (11.8%) include an attribute relating to reproduction (male and female) and two (11.8%) had miscellaneous attributes that fitted into no other category.

Eight studies (47.1%) looked at what mode of administration of DMTs PwMS prefer.[1] All included the options of oral medication and injection, though only 3/8 (37.5%) distinguished between subcutaneous and intramuscular injection, and 5/8 (62.5%) included intravenous infusion.[2] All but one of these studies (Utz, et al. [34]) combined mode and frequency of administration into a single attribute with the disadvantage of making it impossible to fully disentangle their effects. On the other hand, in practice there is a certain amount of correlation between mode and frequency, and combining them a priori rules out unrealistic combinations such as daily intravenous infusions or monthly pills. It also has the advantage of "freeing up" an attribute to describe some other aspect of treatment.

3.7. Probability

Both the benefits and risks of DMTs are probabilistic in nature [10]. The majority of studies investigating preferences with DMTs (11 out of 17, 64.7%) did not explicitly quantify the probability of receiving a given benefit or experiencing a given adverse event. Only a single study (Reed Johnson, et al. [29]) clearly documented using visual means to convey probabilistic information[3], using both a risk grid (a square grid with shaded squares indicating how many patients experience the relevant outcome, e.g. 5 shaded squares out of 1000 to indicate a 0.5%

[1] Although 12 studies had mode of administration as an attribute as an attribute, two studies measured only how important it was in general to participants, without examining preferences between different modes.
[2] Note that although infusion treatments are among the more modern treatments, it is not the case that the three studies that exclude them predate their introduction, coming from 2014, 2015 and 2016.
[3]Wilson, et al. [35] stated that "the visual risk scale was given for reference" but does not elaborate further.

risk) and a risk ladder (a scale giving the context of a given probability in terms of more familiar risks). No study examined how the presentation of probabilities influences preferences.

3.8. Analysis methods

Table 1 lists the main method of analysis for each study. The most popular method was mixed logit, with 10 out of 17 (58.8%) studies, far ahead of the next most popular method, latent class, with three out of 17 (17.6%). To analyse their data, four out of 17 (23.5%) studies used Sawtooth (Sawtooth Software). NLOGIT (Econometric Software) and SPSS (IBM) were each used by three studies (17.6%). Four studies (23.5%) did not report what software they used for analysis.

3.9. Preference heterogeneity according to respondent characteristics

Addressing the needs of individual patients is a crucial part of shared decision making [37], so it is important to go beyond mean preferences to examine how preferences vary according to the characteristics of PwMS. Only eight studies (47.1%) linked respondent heterogeneity to observed characteristics, for example by including them as covariates in a regression [34, 35] or using latent class analysis [26, 36]. However, some others accounted for heterogeneity by using models such as mixed logit without linking it to respondent characteristics. The aspects of heterogeneity considered by each study were categorised by one author (EW). Figure 3 illustrates how many studies examined a given category. Seven out of eight studies (87.5%) tested for the influence of past or current treatment. Several studies explored heterogeneity according to demographic factors (age, gender, education), disease related factors (disease status/history, diagnosis) or quality of life related factors (for example the influence on lifestyle of pain and fatigue).

3.10. ISPOR conjoint analysis checklist quality assessment

In general, all studies scored well against the ISPOR conjoint analysis checklist, with a median count of 23/30 (range 18 – 27). Variation was low both in total counts as well as most counts for individual items. The checklist was not useful in highlighting otherwise unconsidered

issues. Due to this and its limited ability to discriminate between studies, we do not consider its use a success.

4. Discussion

We performed a systematic review of 17 AbSP studies in the field of MS. All but one study investigated the preferences of PwMS for aspects of DMTs, highlighting the importance of trade-offs when considering long term treatment of this chronic condition, which makes DMTs an obvious topic for AbSP techniques.

The most common survey method employed was DCE, which is consistent with a greater number of DCEs than other types of survey in healthcare in general. [20] found only 62 BWS studies in total published up until April 2016, whereas [19] found 179 DCEs between 2009 and 2012 alone). It also reflects that the structure of DCEs, i.e. choosing between two or more alternatives, is closer to the target decision making situation of most studies, choosing between different DMTs, than other study types.

A consequence of the focus on DMTs is the higher proportion of studies being directed only towards PwMS who have a diagnosis of RRMS (42.9%) where there are considerably more licensed DMTs available compared to progressive MS, for which so far only one drug, ocrelizumab (Ocrevus®) has been approved [38, 39]. However, this is still an improvement on the situation a few years ago. Patients with RRMS and PPMS are distinct groups, which differ in terms of past experience and projected disease course. This brings difficulties in capturing preference information from both groups using a single instrument. However, it also means that the preferences of people with different diagnoses may be very different from each other. Thus, while the literature's focus on PwRRMS was appropriate in the past, the anticipated arrival of DMTs for progressive forms of MS means there is now also a need for research into the preferences of PwPPMS.

The paucity of use of qualitative methods to develop attributes thus reflects an area for improvement, as well as better documentation of how attributes are developed and selected. It is not always appropriate to undertake extensive qualitative work in attribute development. (For example, [40] uses the well-known EQ-5D descriptions of health states as attributes, so

qualitative work developing attributes would be nonsensical.) However, PwMS are a heterogeneous population with a large variety of health-related experiences. To avoid omitting crucial aspects of decision making in AbSP studies, it is therefore vital to involve PwMS in attribute development.

PwMS are in most cases non-medical professionals and as their disease progresses many suffer from cognitive impairment. Hence, even if attributes are largely dictated by the research question, qualitative interviews are useful in identifying the best way of meaningfully expressing attributes to participants. For example, a standard measure of the impact of a DMT on individuals' future functioning in clinical trials is the number of people experiencing an increase of 0.5-1 on the expanded disability status scale (EDSS) over a 3-6 month period [41]. Such a measure is difficult to translate into a concept meaningful to PwMS. Qualitative work is thus particularly needed when developing and selecting attributes for AbSP studies in MS.

The majority of studies analysed their data using advanced modelling techniques such as mixed logit. However, several studies used a mixed logit model, but referred to it as a hierarchical Bayes model, or hierarchical Bayes analysis. This nomenclature is incorrect, as hierarchical Bayes is not a model itself, but rather an estimation method used to obtain the parameters of a model [42]. Many studies did not employ analytical techniques that examine response heterogeneity according to observed characteristics.

Given the diverse manifestations and chronic deteriorating nature of MS, it is particularly important to consider response heterogeneity if studies are to accurately reflect the range of patient experiences and opinions. For example, it would be interesting to examine the influence on decision making of risk preference due to the risks associated with DMTs, or cognition, given the cognitive impairments many PwMS experience. Comparison of preferences and priorities of patients at different disease stages would also offer important insights into how experiences impact decisions. It should be noted, however, that including respondent characteristics when analysing AbSP data can be difficult due to the additional model parameters introduced.

Only a single study offered and opt-out option. A significant number of PwMS choose not to take any DMT [43]. Hence, offering only a forced choice between DMTs means not capturing this aspect of their preferences. On the other hand, an opt-out also mean losing some

information about respondents' preferences between different options, as well as risking that people choose the opt-out only to avoid making difficult choices [44]. Thus it is by no means appropriate for every study. However, studies which give only a forced choice should justify this decision and discuss its impact on their results. Future studies should also consider alternative to an opt-out such as a dual response design (respondents first make a forced-choice, then indicate if they would prefer to opt out).

The design of a stated preference survey is crucial for the interpretation of its results [25, 45]. Many studies failed to report the criteria by which they constructed their design, making it impossible for the reader to judge whether it was done appropriately. In addition, different software packages, and different versions of software packages, each have their own algorithms for design construction. Thus, it is important for this to be reported for study reproducibility, which several studies did not.

The studies reviewed employed a wide range of sample sizes, and it is often difficult to assess whether they have recruited appropriate numbers of participants. Several "rules of thumb" for AbSP sample size exist [46, 47] as well as guides for calculation [48]. Thus sample size considerations, whether explicitly calculated using priors or by less formal methods, are possible and usually necessary, and should be both undertaken and reported in future studies. If it is not known whether researchers have achieved an appropriate number of responses, it causes problems for assessing the quality and validity of its results.

Only two studies ([35, 36]) justified the mode of survey administration used, although it should be noted that in many cases the authors may have felt the justification to be self-evident to the reader (e.g. a population drawn from an online community). Nevertheless, given the physical and cognitive impairments experienced by many PwMS which may impact the accessibility of surveys, it would be an improvement for future surveys to document that such factors have been considered. Studies using a convenience sample from a clinic should also show they have considered the impact of this choice on the representativeness of their responses.

The most common attributes were related to prevention of relapses, progression, and side effects, which are probabilistic in nature. Yet the majority of studies presented the outcome of treatment decisions as certainties, e.g. respondents were certain to experience two relapses over the next four years. People's preferences for probabilistic outcomes are extremely

heterogeneous and can have a significant influence over their decision making. Thus, if preferences are elicited only for benefits/costs states as certainties (e.g. "3 relapses in the next 4 years" [49]) it calls into question the external validity of the results for preferences over real DMTs. It is difficult in general to appropriately communicate probabilities (see [50, 51] for overviews of current best practice). It can be even more difficult for participants to understand multiple probabilistic attributes. Thus for pragmatic reasons it is sometimes necessary to represent probabilistic aspects of treatments as certainties. However, given that a majority of studies have no probabilistic representation at all, the appropriate and regular inclusion of probabilistic aspects of DMTs is thus a feature of the literature in need of improvement. In addition, if probabilistic outcomes are represented as deterministic to aid respondent comprehension, the possible impact of this should be discussed.

There is evidence that different ways of presenting probabilities influence individuals' understanding of them [51, 52], and also that understanding can be improved by using graphic presentation of probabilities [53, 54]. However, those studies that did use probabilistic attributes did not report whether they considered their mode of presentation appropriate. Only one study displayed probabilistic information visually using graphs or pictographs. None of the studies explored how choices are influenced by different modes of presentation. Likewise, no study examined the impact on PwMS' DMT preferences over Knightian uncertainty (outcomes whose probabilities of occurring are not explicitly quantified, or "unknown unknowns" [55]), although the long term effects of DMTs are in many cases unknown and even in the short term the risks of rare side effects may not be well quantified [56, 57].

Given so many aspects of MS and DMTs are characterised by poorly quantified risk and uncertainty, there is a need for better tools to communicate risk[1], particularly against the backdrop of the cognitive impairment associated with MS.

Only two studies included an attribute related to reproduction, and in neither of these two did it play a significant role in the analysis. We believe this to be an understudied area, due to previous research highlighting its importance [59, 60], the higher incidence of MS in women of child-bearing age [61], and the variation in advice regarding conception, pregnancy and

[1] It should be noted that this recommendation could apply to AbSP studies in health in general, and not just limited to the field of MS [58].

breastfeeding [62]. In addition, there is a lack of clinical research into the influence of DMTs on reproduction [63, 64] and some DMTs are contraindicated for men with MS trying to conceive [65].

The vast majority of studies (88.2%) were funded by pharmaceutical companies. These companies have many reasons to fund studies, for example information about patient preferences can be useful both in marketing existing drugs and informing future drug development. It can also be used to aid regulatory decisions, for example with the US Federal Drug Administration[1]. Pharmaceutical companies are the funders of nearly all the AbSp studies in MS. Thus nearly all of the literature addresses the aims of pharmaceutical companies. It would be welcome to have a broader range of funders, as this would bring a broader range of research objectives and greater diversity of studies.

A strength of our work is its focus on the technical aspects of AbSP studies in MS. The number of such studies is increasing over time, and it can serve as a guide to the details of running them, and be a practical aid to future research.

Another strength is that we have highlighted several areas of current practice which can be improved, particularly greater use of qualitative methods and better reporting of survey design choices. We have also highlighted gaps in the current literature. Future studies may wish to consider examining patient preferences surrounding DMTs and reproduction, how different methods of risk communication affect the decision making of PwMS, or the effect of Knightian uncertainty.

Our study has several limitations. We have not quantitatively combined the results of studies. We took this decision partly because of our focus on study design, and also due to the difficulties of combining numerical results from studies using different methodologies and different ways of presenting results, as well as different attributes and level sets. Nevertheless, in the future, it would be informative to attempt a synthesis of results from AbSP studies in MS.

[1] See e.g. https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm446680.pdf.

Assessing study quality using the ISPOR conjoint analysis checklist is not recommended for future reviews of AbSP studies. It did not distinguish between minimum acceptable practice, for example basing attributes on a non-systematic review of clinical literature, and good practice, for example developing attributes through extensive qualitative research. That it was not a good measure of quality is perhaps unsurprising, as it was not created for that purpose, but rather as a rough guide to best practice when developing surveys [25].

Due to the focus on details of study design, unpublished studies such as conference proceedings were excluded, as we felt they included insufficient methodological information.

5. Conclusion

Shared decision making including patient preference views on treatment are increasingly used in medicine, particularly chronic conditions, such as MS, with a still evolving DMT landscape [66, 67]. Thus, it is important to investigate patient preferences, especially when the experiences of PwMS are very heterogeneous and there are many treatment options available.

AbSP studies such as DCEs are increasingly used to measure PwMS' preferences providing insights into this field. Several areas in need of improvement have been highlighted, particularly a greater use of qualitative methods in attribute and level development. Further work should be undertaken to better characterise the role of reproduction in decision making, and better communication of risk is warranted.

Data availability statement: Data extracted from studies is included as supplementary material.

References

[1] World Health Organization, "Atlas: multiple sclerosis resources in the world 2008," 2008.

[2] W. J. Brownlee, T. A. Hardy, F. Fazekas, and D. H. Miller, "Diagnosis of multiple sclerosis: progress and challenges," The Lancet, vol. 389, pp. 1336-1346.

[3] A. Compston and A. Coles, "Multiple sclerosis," The Lancet, vol. 372, pp. 1502-1517.

[4] C. Stadelmann, C. Wegner, and W. Brück, "Inflammation, demyelination, and degeneration—recent insights from MS pathology," Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease, vol. 1812, pp. 275-282, 2011.

[5] E. Leray, J. Yaouanq, E. Le Page, M. Coustans, D. Laplaud, J. Oger, et al., "Evidence for a two-stage disability progression in multiple sclerosis," Brain, vol. 133, pp. 1900-1913, 2010.

[6] A. Scalfari, A. Neuhaus, A. Degenhardt, G. P. Rice, P. A. Muraro, M. Daumer, et al., "The natural history of multiple sclerosis, a geographically based study 10: relapses and long-term disability," Brain, vol. 133, pp. 1914-1929, 2010.

[7] M. Rovaris, C. Confavreux, R. Furlan, L. Kappos, G. Comi, and M. Filippi, "Secondary progressive multiple sclerosis: current knowledge and future challenges," The Lancet Neurology, vol. 5, pp. 343-354, 2006.

[8] H. Lassmann, J. van Horssen, and D. Mahad, "Progressive multiple sclerosis: pathology and pathogenesis," Nat Rev Neurol, vol. 8, pp. 647-656, 11//print 2012.

[9] G. Giovannoni, J. A. Cohen, A. J. Coles, H.-P. Hartung, E. Havrdova, K. W. Selmaj, et al., "Alemtuzumab improves preexisting disability in active relapsing-remitting MS patients," Neurology, vol. 87, pp. 1985-1992, November 8 2016.

[10] G. Comi, M. Radaelli, and P. Soelberg Sørensen, "Evolving concepts in the treatment of relapsing multiple sclerosis," The Lancet, vol. 389, pp. 1347-1356.

[11] M. Ryan, K. Gerard, and M. Amaya-Amaya, Using discrete choice experiments to value health and health care vol. 11: Springer Science & Business Media, 2007.

[12] S. Simioni, M. Schluep, N. Bault, G. Coricelli, J. Kleeberg, R. A. Du Pasquier, et al., "Multiple sclerosis decreases explicit counterfactual processing and risk taking in decision making," PloS one, vol. 7, p. e50718, 2012.

[13] N. Muhlert, V. Sethi, L. Cipolotti, H. Haroon, G. J. Parker, T. Yousry, et al., "The grey matter correlates of impaired decision-making in multiple sclerosis," Journal of Neurology, Neurosurgery & Psychiatry, vol. 86, pp. 530-536, 2015.

[14] M. G. Cogo, S. Rota, M. L. Fusco, C. Mapelli, F. Ferri, I. M. Appollonio, et al., "Cognitive correlates of under-ambiguity and under-risk decision making in high-functioning patients with relapsing remitting multiple sclerosis," Journal of clinical and experimental neuropsychology, vol. 36, pp. 1066-1075, 2014.

[15] M. F. Farez, L. Crivelli, R. Leiguarda, and J. Correale, "Decision-making impairment in patients with multiple sclerosis: a case–control study," BMJ open, vol. 4, p. e004918, 2014.

[16] A. D. Radomski, C. Power, S. E. Purdon, D. J. Emery, G. Blevins, K. G. Warren, et al., "Decision-making under explicit risk is impaired in multiple sclerosis: relationships with ventricular width and disease disability," BMC neurology, vol. 15, p. 61, 2015.

[17] M. Ryan and K. Gerard, "Using discrete choice experiments to value health care programmes: current practice and future research reflections," Appl Health Econ Health Policy, vol. 2, pp. 55-64, 2003.

[18] E. W. de Bekker-Grob, M. Ryan, and K. Gerard, "Discrete choice experiments in health economics: a review of the literature," Health Economics, vol. 21, pp. 145-172, 2012.

[19] M. D. Clark, D. Determann, S. Petrou, D. Moro, and E. W. de Bekker-Grob, "Discrete Choice Experiments in Health Economics: A Review of the Literature," PharmacoEconomics, vol. 32, pp. 883-902, 2014.

[20] K. L. Cheung, B. F. Wijnen, I. L. Hollin, E. M. Janssen, J. F. Bridges, S. M. Evers, et al., "Using Best-Worst Scaling to Investigate Preferences in Health Care," Pharmacoeconomics, vol. 34, pp. 1195-1209, Dec 2016.

[21] J. McGowan, M. Sampson, D. M. Salzwedel, E. Cogo, V. Foerster, and C. Lefebvre, "PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Statement," J Clin Epidemiol, vol. 75, pp. 40-6, Jul 2016.

[22] M. Sampson, J. McGowan, E. Cogo, J. Grimshaw, D. Moher, and C. Lefebvre, "An evidence-based practice guideline for the peer review of electronic search strategies," J Clin Epidemiol, vol. 62, pp. 944-52, Sep 2009.

[23] J. Popay, H. Roberts, A. Sowden, M. Petticrew, L. Arai, M. Rodgers, et al., "Guidance on the conduct of narrative synthesis in systematic reviews," A product from the ESRC methods programme Version, vol. 1, p. b92, 2006.

[24] J. J. Louviere and E. Lancsar, "Choice experiments in health: the good, the bad, the ugly and toward a brighter future," Health Economics, Policy and Law, vol. 4, pp. 527-546, 2009.

[25] J. F. Bridges, A. B. Hauber, D. Marshall, A. Lloyd, L. A. Prosser, D. A. Regier, et al., "Conjoint analysis applications in health—a checklist: a report of the ISPOR Good Research Practices for Conjoint Analysis Task Force," Value in Health, vol. 14, pp. 403-413, 2011.

[26] R. Rosato, S. Testa, A. Oggero, G. Molinengo, and A. Bertolotto, "Quality of life and patient preferences: identification of subgroups of multiple sclerosis patients," Quality of Life Research, vol. 24, pp. 2173-82, Sep 2015.

[27] I. E. Kremer, S. M. Evers, P. J. Jongen, T. van der Weijden, I. van de Kolk, and M. Hiligsmann, "Identification and Prioritization of Important Attributes of Disease-Modifying Drugs in Decision Making among Patients with Multiple Sclerosis: A Nominal Group Technique and Best-Worst Scaling," PLoS ONE [Electronic Resource], vol. 11, p. e0164862, 2016.

[28] L. D. Lynd, A. Traboulsee, C. A. Marra, N. Mittmann, C. Evans, K. H. Li, et al., "Quantitative analysis of multiple sclerosis patients' preferences for drug treatment: a best-worst scaling study," Therapeutic Advances in Neurological Disorders, vol. 9, pp. 287-96, Jul 2016.

[29] F. Reed Johnson, G. Van Houtven, S. Ozdemir, S. Hass, J. White, G. Francis, et al., "Multiple sclerosis patients' benefit-risk preferences: serious adverse event risks versus treatment efficacy," Journal of Neurology, vol. 256, pp. 554-62, Apr 2009.

[30] C. Poulos, E. Kinter, J. C. Yang, J. F. Bridges, J. Posner, E. Gleisner, et al., "A discrete-choice experiment to determine patient preferences for injectable multiple sclerosis treatments in Germany," Therapeutic Advances in Neurological Disorders, vol. 9, pp. 95-104, Mar 2016.

[31] P. Wicks, D. Brandes, J. Park, D. Liakhovitski, T. Koudinova, and R. Sasane, "Preferred features of oral treatments and predictors of non-adherence: two web-based choice experiments in multiple sclerosis patients," Interactive Journal of Medical Research, vol. 4, p. e6, Mar 05 2015.

[32] D. Marshall, J. F. Bridges, B. Hauber, R. Cameron, L. Donnalley, K. Fyie, et al., "Conjoint analysis applications in health—how are studies being designed and reported?," The Patient: Patient-Centered Outcomes Research, vol. 3, pp. 249-256, 2010.

[33] F. Reed Johnson, E. Lancsar, D. Marshall, V. Kilambi, A. Muhlbacher, D. A. Regier, et al., "Constructing experimental designs for discrete-choice experiments: report of the ISPOR Conjoint Analysis Experimental Design Good Research Practices Task Force," Value Health, vol. 16, pp. 3-13, Jan-Feb 2013.

[34] K. S. Utz, J. Hoog, A. Wentrup, S. Berg, A. Lammer, B. Jainsch, et al., "Patient preferences for disease-modifying drugs in multiple sclerosis therapy: a choice-based conjoint analysis," Therapeutic Advances in Neurological Disorders, vol. 7, pp. 263-75, Nov 2014.

[35] L. S. Wilson, A. Loucks, G. Gipson, L. Zhong, C. Bui, E. Miller, et al., "Patient preferences for attributes of multiple sclerosis disease-modifying therapies: development and results of a ratings-based conjoint analysis," International Journal of Ms Care, vol. 17, pp. 74-82, Mar-Apr 2015.

[36] C. Bottomley, A. Lloyd, G. Bennett, and N. Adlard, "A discrete choice experiment to determine UK patient preference for attributes of disease modifying treatments in Multiple Sclerosis," Journal of Medical Economics, pp. 1-8, 2017.

[37] A. M. Stiggelbout, T. Van der Weijden, M. P. De Wit, D. Frosch, F. Légaré, V. M. Montori, et al., "Shared decision making: really putting patients at the centre of healthcare," BMJ: British Medical Journal (Online), vol. 344, 2012.

[38] S. L. Hauser, A. Bar-Or, G. Comi, G. Giovannoni, H.-P. Hartung, B. Hemmer, et al., "Ocrelizumab versus Interferon Beta-1a in Relapsing Multiple Sclerosis," New England Journal of Medicine, vol. 376, pp. 221-234, 2017.

[39] X. Montalban, S. L. Hauser, L. Kappos, D. L. Arnold, A. Bar-Or, G. Comi, et al., "Ocrelizumab versus Placebo in Primary Progressive Multiple Sclerosis," New England Journal of Medicine, vol. 376, pp. 209-220, 2017.

[40] M. F. Jonker, A. E. Attema, B. Donkers, E. A. Stolk, and M. M. Versteegh, "Are Health State Valuations from the General Public Biased? A Test of Health State Reference Dependency Using Self-assessed Health and an Efficient Discrete Choice Experiment," Health economics, 2016.

[41] A. M. Lavery, L. H. Verhey, and A. T. Waldman, "Outcome measures in relapsing-remitting multiple sclerosis: capturing disability and disease progression in clinical trials," Multiple sclerosis international, vol. 2014, 2014.

[42] K. E. Train, "Discrete choice methods with simulation," ed: Cambridge university press, 2009, pp. 282-314.

[43] N. Grytten, J. Aarseth, K. Espeset, G. Berg Johnsen, R. Wehus, C. Lund, et al., "Health-related quality of life and disease-modifying treatment behaviour in relapsing-remitting multiple sclerosis–a multicentre cohort study," Acta Neurologica Scandinavica, vol. 126, pp. 51-57, 2012.

[44] J. Veldwijk, M. S. Lambooij, E. W. de Bekker-Grob, H. A. Smit, and G. A. De Wit, "The effect of including an opt-out option in discrete choice experiments," PloS one, vol. 9, p. e111805, 2014.

[45] J. M. Rose and M. C. J. Bliemer, "Sample size requirements for stated choice experiments," Transportation, vol. 40, pp. 1021-1041, 2013.

[46] R. Johnson and B. Orme, "Getting the most from CBC," Sequim: Sawtooth Software Research Paper Series, Sawtooth Software, 2003.

[47] E. Lancsar and J. Louviere, "Conducting Discrete Choice Experiments to Inform Healthcare Decision Making," PharmacoEconomics, vol. 26, pp. 661-677, 2008.

[48] E. W. de Bekker-Grob, B. Donkers, M. F. Jonker, and E. A. Stolk, "Sample Size Requirements for Discrete-Choice Experiments in Healthcare: a Practical Guide," The Patient - Patient-Centered Outcomes Research, vol. 8, pp. 373-384, 2015.

[49] C. Poulos, E. Kinter, J. C. Yang, J. F. Bridges, J. Posner, and A. T. Reder, "Patient Preferences for Injectable Treatments for Multiple Sclerosis in the United States: A Discrete-Choice Experiment," The Patient: Patient-Centered Outcomes Research, vol. 9, pp. 171-80, Apr 2016.

[50] D. Spiegelhalter, "Risk and Uncertainty Communication," Annual Review of Statistics and Its Application, vol. 4, pp. 31-60, 2017.

[51] A. J. Apter, M. K. Paasche-Orlow, J. T. Remillard, I. M. Bennett, E. P. Ben-Joseph, R. M. Batista, et al., "Numeracy and communication with patients: they are counting on us," J Gen Intern Med, vol. 23, pp. 2117-24, Dec 2008.

[52] D. G. Goldstein and D. Rothschild, "Lay understanding of probability distributions," Judgment and Decision Making, vol. 9, p. 1, 2014.

[53] R. Hess, V. H. Visschers, and M. Siegrist, "Risk communication with pictographs: the role of numeracy and graph processing," Judgment and Decision Making, vol. 6, p. 263, 2011.

[54] R. Garcia-Retamero, Y. Okan, and E. T. Cokely, "Using visual aids to improve communication of risks about health: a review," The Scientific World Journal, vol. 2012, 2012.

[55] F. H. Knight, Risk, Uncertainty and Profit. Boston, MA: Hart, Schaffner & Marx; Houghton Mifflin Co., 1921.

[56] W. Brück, R. Gold, B. T. Lund, C. Oreja-Guevara, A. Prat, C. M. Spencer, et al., "Therapeutic decisions in multiple sclerosis: moving beyond efficacy," JAMA neurology, vol. 70, pp. 1315-1324, 2013.

[57] W. Castro-Borrero, D. Graves, T. C. Frohman, A. B. Flores, P. Hardeman, D. Logan, et al., "Current and emerging therapies in multiple sclerosis: a systematic review," Therapeutic Advances in Neurological Disorders, vol. 5, pp. 205-220, 2012.

[58] M. Harrison, D. Rigby, C. Vass, T. Flynn, J. Louviere, and K. Payne, "Risk as an attribute in discrete choice experiments: a systematic review of the literature," The Patient-Patient-Centered Outcomes Research, vol. 7, pp. 151-170, 2014.

[59] A. Carvalho, A. Veiga, J. Morgado, R. Tojal, S. Rocha, J. Vale, et al., "Multiple sclerosis and motherhood choice: an observational study in Portuguese women patients," Revista de neurologia, vol. 59, pp. 537-542, 2014.

[60] D. Payne and K. M. McPherson, "Becoming mothers. Multiple sclerosis and motherhood: a qualitative study," Disability and rehabilitation, vol. 32, pp. 629-638, 2010.

[61] H. F. Harbo, R. Gold, and M. Tintoré, "Sex and gender issues in multiple sclerosis," Therapeutic advances in neurological disorders, vol. 6, pp. 237-248, 2013.

[62] P. K. Coyle, "Management of women with multiple sclerosis through pregnancy and after childbirth," Therapeutic Advances in Neurological Disorders, vol. 9, pp. 198-210, 03/02 2016.

[63] R. Alroughani, A. Altintas, M. Al Jumah, M. Sahraian, I. Alsharoqi, A. AlTahan, et al., "Pregnancy and the Use of Disease-Modifying Therapies in Patients with Multiple Sclerosis: Benefits versus Risks," Mult Scler Int, vol. 2016, p. 1034912, 2016.

[64] D. H. Miller, F. Fazekas, X. Montalban, S. C. Reingold, and M. Trojano, "Pregnancy, sex and hormonal factors in multiple sclerosis," Multiple Sclerosis Journal, vol. 20, pp. 527-536, 2014.

[65] R. Bove, A. McHenry, K. Hellwig, M. Houtchens, N. Razaz, P. Smyth, et al., "Multiple sclerosis in men: management considerations," Journal of neurology, vol. 263, pp. 1263-1273, 2016.

[66] S. Köpke, S. Kern, T. Ziemssen, M. Berghoff, I. Kleiter, M. Marziniak, et al., "Evidence-based patient information programme in early multiple sclerosis: a randomised controlled trial," J Neurol Neurosurg Psychiatry, vol. 85, pp. 411-418, 2014.

[67] J. Kasper, S. Köpke, I. Mühlhauser, M. Nübling, and C. Heesen, "Informed shared decision making about immunotherapy for patients with multiple sclerosis (ISDIMS): a randomized controlled trial," European Journal of Neurology, vol. 15, pp. 1345-1352, 2008.

[68] R. Arroyo, A. P. Sempere, E. Ruiz-Beato, D. Prefasi, A. Carreno, M. Roset, et al., "Conjoint analysis to understand preferences of patients with multiple sclerosis for disease-modifying therapy attributes in Spain: a cross-sectional observational study," BMJ Open, vol. 7, p. e014433, 2017.

[69] C. S. Carlin, L. Higuera, and S. Anderson, "Improving Patient-Centered Care by Assessing Patient Preferences for Multiple Sclerosis Disease-Modifying Agents: A Stated-Choice Experiment," Permanente Journal, vol. 21, 2017.

[70] J. M. Garcia-Dominguez, D. Munoz, M. Comellas, I. Gonzalbo, L. Lizan, and C. Polanco Sanchez, "Patient preferences for treatment of multiple sclerosis with disease-

modifying therapies: a discrete choice experiment," Patient preference & adherence, vol. 10, pp. 1945-1956, 2016.

[71] A. P. Sempere, V. Vera-Lopez, J. Gimenez-Martinez, E. Ruiz-Beato, J. Cuervo, and J. Maurino, "Using a multidimensional unfolding approach to assess multiple sclerosis patient preferences for disease-modifying therapy: A pilot study," Patient Preference and Adherence, vol. 11, pp. 995-999, 29 May 2017.

[72] S. L. Shingler, P. Swinburn, S. Ali, R. Perard, and A. J. Lloyd, "A discrete choice experiment to determine patient preferences for injection devices in multiple sclerosis," Journal of Medical Economics, vol. 16, pp. 1036-42, Aug 2013.

[73] L. Wilson, A. Loucks, C. Bui, G. Gipson, L. Zhong, A. Schwartzburg, et al., "Patient centered decision making: use of conjoint analysis to determine risk-benefit trade-offs for preference sensitive treatment choices," Journal of the Neurological Sciences, vol. 344, pp. 80-7, Sep 15 2014.

Figure 1: PRISMA flow diagram

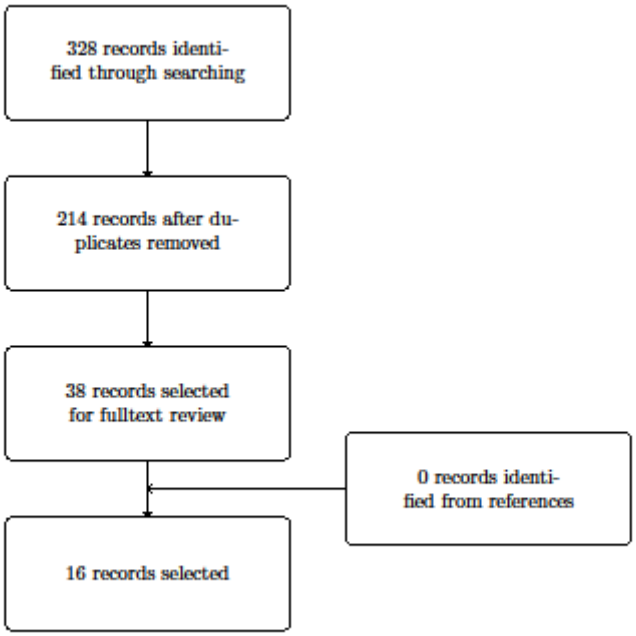

Figure 2: Number of publications of attribute-based stated preference studies in multiple sclerosis per year

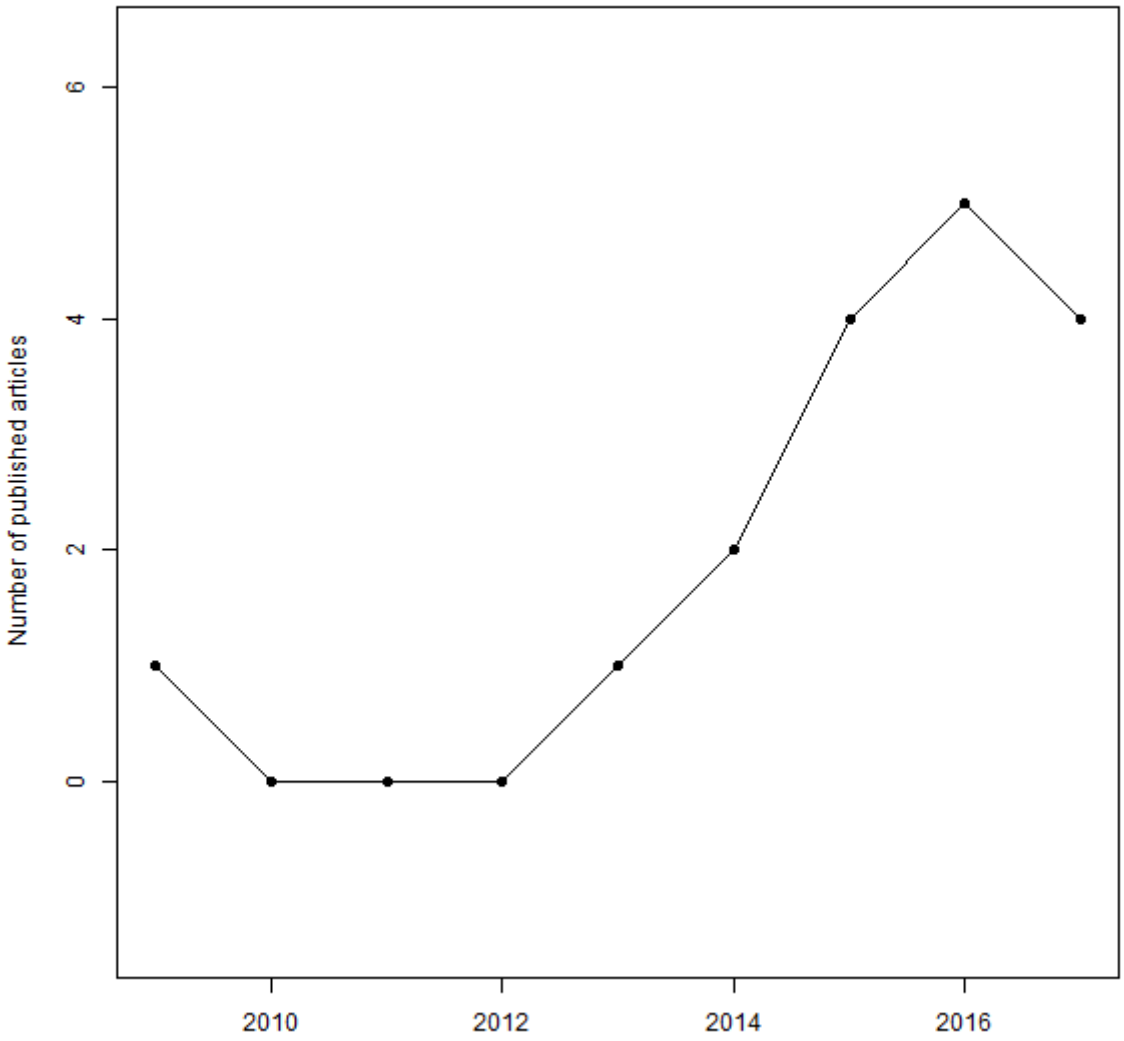

Figure 3: (a) Number of attributes included in each study; (b) number of questions answered per subject in each study; (c) sample sizes for final analysis for each study; (d) number of studies examining a given aspect of preference heterogeneity.

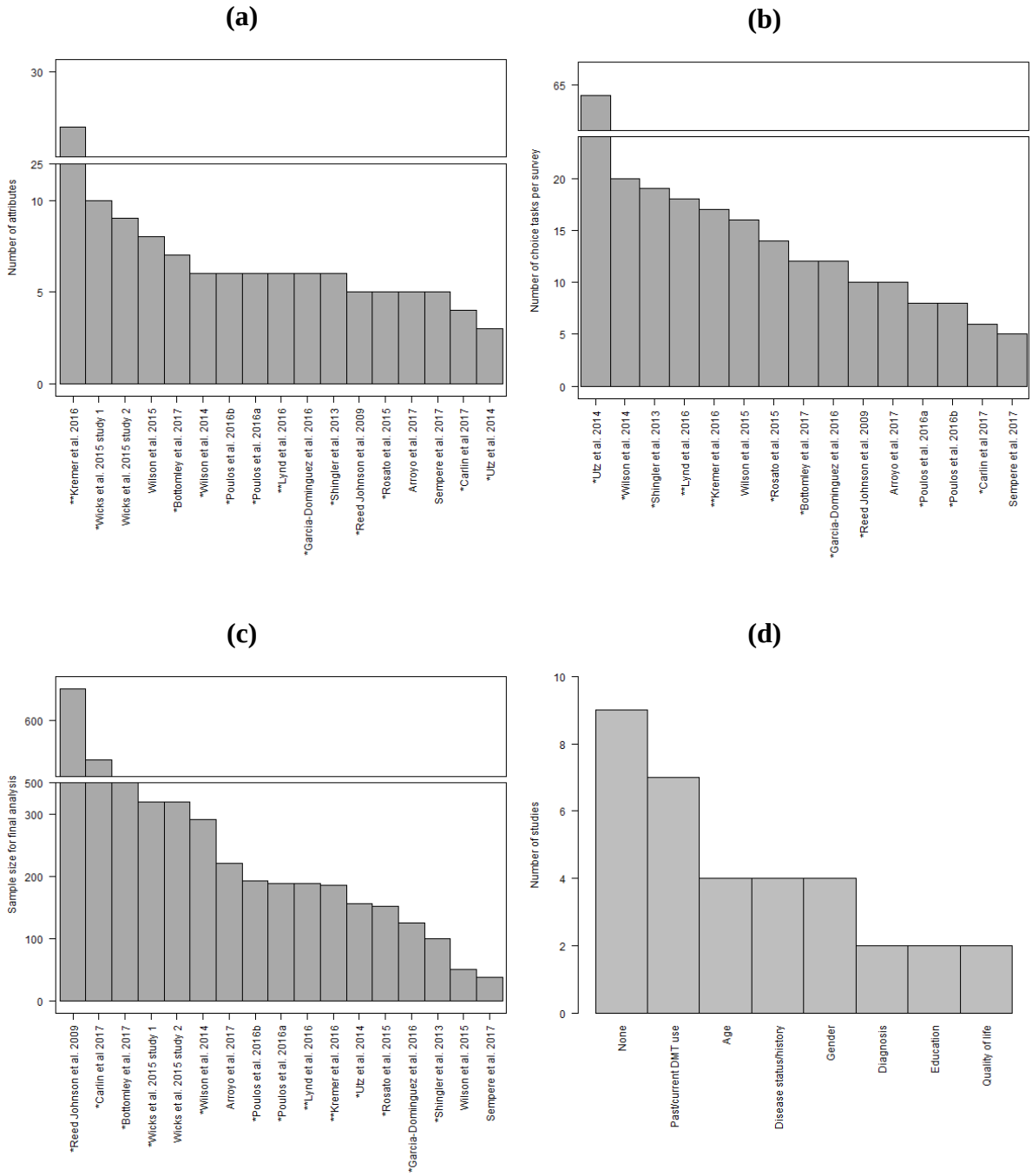

Table 1: Studies included for review. ES = Spain, NL = Netherlands, CA = Canada, DE = Germany, IT = Italy, DMT = disease modifying treatment, DCE = discrete choice experiment, BWS = best-worst scaling, MDU = multidimensional unfolding, LR = linear regression, LC = latent class, MNP = multinomial probit, MIX = mixed logit, SMR = stepwise multilinear regression, CL = conditional logit.

Study	Country	Topic	Type	Analysis	Sample Size	No. of attributes	Mean levels per attribute	Questions per survey
Arroyo et al. (2017) [68]	ES	DMTs	Rating	LR	221	5	2.6	10
Bottomley et al. (2017) [36]	UK	DMTs	DCE	LC	350	7	3.29	12
Carlin et al. (2017) [69]	US	DMTs	DCE	MNP	537	4	5	6
Garcia-Dominguez et al. (2016) [70]	ES	DMTs	DCE	MIX, SMR	125	6	2.67	12
Kremer et al. (2016) [27]	NL	DMTs	BWS (case 1)	MIX	185	27	1	17
Lynd et al. (2016) [28]	CA	DMTs	BWS (case 2)	CL, LC	189	6	3.17	18
Poulos et al. (2016a) [30]	DE	DMTs	DCE	MIX	192	6	3.17	8
Poulos et al. (2016b) [49]	US	DMTs	DCE	MIX	189	6	3.17	8
Reed Johnson et al. (2009) [29]	US	DMTs	DCE	MIX	651	5	4	10
Rosato et al. (2015) [26]	IT	Quality of life	DCE	LC	152	5	3	14
Sempere et al. (2017) [71]	ES	DMTs	MDU	MDU	37	5	1	5
Shingler et al. (2013) [72]	UK	DMTs	DCE	MIX	99	6	3	19
Utz et al. (2014) [34]	DE	DMTs	DCE	MIX	156	3	2.67	64
Wicks et al. (2015) [31] study 1	US	DMTs	DCE	MIX	319	10	2.9	Not stated
Wicks et al. (2015) [31] study 2	US	DMTs	Rating	MIX	319	9	2.11	Not stated
Wilson et al. (2014) [73]	US	DMTs	DCE	MIX	291	6	3.33	20

Wilson et al. (2015) [35]	US	DMTs	Ranking	MME	50	8	4	16

Appendix A Search strategy

CINAHL (EBSCO) 1981- present

Print Search History Monday, July 10, 2017 5:00:08 PM
S15 (S6 AND S14) 17
S14 S7 OR S8 OR S9 OR S10 OR S11 OR S12 OR S13 3,286
S13 TI preference* based OR AB preference* based 1,027
S12 TI ((choice N1 (based or model* or experiment* or behavio#r*))) OR AB ((choice N1 (based or model* or experiment* or behavio#r*))) 1,018
S11 TI ((maximum difference or maxdiff or max-diff)) OR AB ((maximum difference or maxdiff or max-diff)) 664
S10 TI ((BWS or best-worst)) OR AB ((BWS or best-worst).) 178
S9 TI conjoint OR AB conjoint 345
S8 TI stated preference*. OR AB stated preference*. 161
S7 TI discrete choice*. OR AB discrete choice* 313
S6 S1 OR S2 OR S3 OR S4 OR S5 12,050
S5 (MH "Myelitis, Transverse+") 138
S4 TI transverse myelitis OR AB transverse myelitis 188
S3 TI encephalomyelitis disseminata OR AB encephalomyelitis disseminata 1
S2 TI (((multiple or disseminated) N2 scleros*)) OR AB (((multiple or disseminated) N2 scleros*)) 7,716
S1 (MH "Multiple Sclerosis+") 10,585

COCHRANE

Date Run: 10/07/17 21:56:12.926
Description: 08-02-17

- Cochrane Database of Systematic Reviews (Wiley) : Issue 7 of 12, July 2017 (n = 2)
- Cochrane Central Register of Controlled Trials (Wiley) : Issue 6 of 12, June 2017 (n = 47)
- Cochrane Methodology Register (Wiley) : Issue 3 of 4, July 2012 (n = 1)

ID	Search	Hits
#1	MeSH descriptor: [Multiple Sclerosis] explode all trees	2344
#2	(multiple scleros* or disseminat* scleros*):ti,ab	5931
#3	encephalomyelitis disseminata:ti,ab	0
#4	Transverse Myelitis:ti,ab	29
#5	MeSH descriptor: [Myelitis, Transverse] explode all trees	15
#6	#1 or #2 or #3 or #4 or #5	6141
#7	(discrete choice):ti,ab	134
#8	(stated preference*):ti,ab	206
#9	MeSH descriptor: [Choice Behavior] explode all trees	1192
#10	conjoint:ti,ab	168
#11	(BWS or best-worst):ti,ab	63
#12	(maximum difference or maxdiff or max-diff):ti,ab	6416
#13	(choice near/1 (based or model* or experiment* or behavio?r*)):ti,ab	203
#14	(preference* based):ti,ab	2171
#15	#7 or #8 or #9 or #10 or #11 or #12 or #13 or #14	10203

#16 #6 and #15 50

Database: Embase Classic+Embase <1947 to 2017 July 07>

---

1 multiple sclerosis/ (109982)
2 ((multiple or disseminated) adj2 scleros*).ti,ab. (96116)
3 encephalomyelitis disseminata.ti,ab. (61)
4 transverse myelitis.ti,ab. (3088)
5 myelitis/ (8456)
6 or/1-5 [MS] (125799)
7 discrete choice*.tw. (1795)
8 "discrete choice experiment"/ (152)
9 stated preference*.tw. (598)
10 (BWS or best-worst).tw. (1472)
11 (maximum difference or maxdiff or max-diff).tw. (1514)
12 conjoint.tw. (3003)
13 *patient preference/ (3035)
14 (choice adj (based or model* or experiment* or behavio?r*)).tw. (4926)
15 preference* based.tw. (1598)
16 or/7-15 [DCE terms] (15250)
17 6 and 16 (100)

Database: Ovid MEDLINE(R) Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) <1946 to 10 July 2017 22:32>

Search Strategy:

---

1 exp Multiple Sclerosis/ (52589)
2 ((multiple or disseminated) adj2 scleros*).ti,ab. (64733)
3 encephalomyelitis disseminata.ti,ab. (45)
4 Myelitis, Transverse/ (1278)
5 transverse myelitis.ti,ab. (1784)
6 or/1-5 [MS] (73795)
7 discrete choice*.tw. (1331)
8 stated preference*.tw. (474)
9 Choice Behavior/ (28072)
10 conjoint.tw. (2295)
11 (BWS or best-worst).tw. (1118)
12 (maximum difference or maxdiff or max-diff).tw. (1191)
13 (choice adj (based or model* or experiment* or behavio?r*)).tw. (4310)
14 preference* based.tw. (1217)
15 or/7-14 [discrete choice terms] (36717)
16 6 and 15 (74)

Database: PsycINFO <1806 to July Week 1 2017>

---

1 multiple sclerosis/ (11174)
2 ((multiple or disseminated) adj2 scleros*).ti,ab. (13441)
3 encephalomyelitis disseminata.ti,ab. (5)
4 transverse myelitis.ti,ab. (268)
5 myelitis/ (484)
6 or/1-5 [MS] (14257)
7 discrete choice*.tw. (758)
8 stated preference*.tw. (440)
9 choice behavior/ (16685)
10 (BWS or best-worst).tw. (259)
11 (maximum difference or maxdiff or max-diff).tw. (79)
12 (choice adj (based or model* or experiment* or behavio?r*)).tw. (7104)
13 preference* based.tw. (774)
14 preferences/ (16011)
15 conjoint.tw. (3024)
16 or/7-15 [DCE] (38306)
17 6 and 16 (19)

Web of Science Core Content (Clarivate Analytics) [11 July 2017]

Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, ESCI Timespan=1900-2017

14 68 #13 AND #5
13 42,767 #12 OR #11 OR #10 OR #9 OR #8 OR #7 OR #6
12 2,672 TOPIC: ("preference* based")
11 27,848 TOPIC: ((choice NEAR/1 (based or model* or experiment* or behavio$r*)).)
10 2,388 TOPIC: (("maximum difference" or maxdiff or max-diff))
9 2,582 TOPIC: ((BWS or best-worst))
8 6,444 TOPIC: (conjoint)
7 2,837 TOPIC: ("stated preference*")
6 5,261 TOPIC: ("discrete choice")
5 102,838 #4 OR #3 OR #2 OR #1
4 2,012 TOPIC: ("transverse myelitis")
3 46 TOPIC: ("encephalomyelitis disseminata")
2 101,604 TOPIC: (((multiple or disseminated) NEAR/2 scleros*))
1 100,909 TOPIC: ("multiple sclerosis")

Appendix B Data extraction form

1. Study context

i. What is the type of study?
ii. What topic is the focus of the study?
iii. In what country/countries is the target population?
iv. What is the diagnosis of the target population?
v. Source of funding
vi. Date of publication

2. Participants

i. How were patients contacted?
ii. Criteria for inclusion
iii. Initial sample size after screening
iv. Number of participants completing at least part of the survey
v. Number of participants failing to complete survey
vi. Number of participants excluded for data quality
vii. Criteria for inclusion in final analysis
viii. Number of participants included in final analysis

3. Attribute development

i. Drawn from literature review?
ii. Consultation with medical professionals?
iii. Focus group?
iv. If yes, what was size of focus group?
v. Interviews?
vi. If yes, how many interviews?
vii. Tested using a pilot study?
viii. If yes, how many participated?
ix. Any other notable characteristics of attribute development

4. Attributes

i. Number of attributes
ii. Number of levels for each attribute
iii. List of attributes
iv. List of levels for each attribute
v. Method of risk presentation

5. Survey design

i. Number of options presented per decision
ii. Number attributes presented for each option
iii. Was a no treatment option included?
iv. Anything else asked?

v. If so, what?
vi. Dominant choice question included?
vii. Total number of questions (not including dominant choices)
viii. Number of questions per survey
ix. Number of different surveys
x. Type of design (full factorial, partial factorial, etc.)
xi. How was design constructed? (Sawtooth, SAS, etc.)
xii. What criteria were used to judge the design?
xiii. Were any power/sample size calculations carried out?
xiv. If no, was there any non-quantitive consideration of power/sample size?
xv. Administration method

6. Additional data collected

i. Demographic information collected?
ii. Disease history collected?
iii. Experience with DMTs collected?
iv. Any other data collected?
v. If so, what?

7. Analysis

i. What was the main method of analysis?
ii. Any secondary method(s) of analysis?
iii. Software used for analysis
iv. Were results presented using importance scores?
v. Were results presented using time til progression as numeraire?
vi. Were results presented using money as numeraire (i.e. WTP)?
vii. Were results presented using maximum acceptable risk?
viii. Were results presented using utilities?
ix. Was participant heterogeneity examined in any other way?
x. If so, what was examined?

8. Conclusions

i. What was the most valued positive attribute?
ii. What was the second most valued positive attribute?
iii. What was the most valued negative attribute?
iv. What was the second most valued negative attribute?
v. If looking at mode of administration, what was the most preferred method?
vi. If looking at mode of administration, what was the least preferred method?
vii. Summary of other findings