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.

[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, University of Leeds
[8] Blizard Institute (Neuroscience), Barts and The London School of Medicine & Dentistry, Queen Mary University of London, London, UK
[9] Emergency Care and Acute Medicine Clinical Academic Group Neuroscience, The Royal London Hospital, Barts Health NHS Trust, London, UK
[10] School of Sociology and Social Policy, University of Leeds

Methods: Published articles using an attribute-based stated preference methodology were identified via a systematic search of the literature, and analysed using the ISPOR conjoint analysis checklist.

Results: We identified 16 relevant articles reporting 17 separate studies, all but one focussing on DMT. 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) 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.

Conclusion: Attribute-based stated preference is a useful method to examine preferences of people with Multiple Sclerosis 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, design choices, using qualitative methods for attribute/level development and better reporting of experimental procedures.
- Effect of treatment on reproduction and the influence of risk perception were identifies as understudied topics.

Acknowledgements: This work was partly funded by the UK Multiple Sclerosis Society Grant no. 30, "CRIMSON - Considering Risk and benefits In Multiple Sclerosis treatment selectiON". JC acknowledges the UK National Institute for Health Research (NIHR) University College London Hospitals Biomedical Research Centres Funding scheme and University College London (UCL)

Conflicts of interest: 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 and Apitope.

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 symptoms PwMS experience and 2) Disease modifying treatments (DMTs) intended to alter the natural course of MS with fewer relapses and slowing of disability progression [9]. DMTs require a long term commitment by PwMS, in most cases requiring regular administration of tablets, injections or infusions.

Currently, 13 DMTs are available for the treatment of people with relapsing-remitting MS (PwRMS) whilst only one has been approved (in the USA; as yet none in Europe) for people with progressive forms of MS. The drugs used vary in efficacy in reducing the relapse rate and mode of delivery. Potential side effects range from mild to severe and also differ in frequency [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. Currently, the most effective DMTs are also associated with the highest risk of adverse effects, including rare 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, and how it will progress and its 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], 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 number of AbSP studies in PwMS has risen recently, with almost all examining DMT choice. Although general reviews of DCEs and BWS in health exist [12-15], none has specifically examined MS. We conducted a systematic review of AbSP in PwMS with a focus on the technical aspects of design and conduct of experiments.

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 using the concepts multiple sclerosis and discrete choice experiments. 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

[1] Note that conjoint analysis is sometimes used as an umbrella term to refer to all studies which measure trade-offs between attributes. However, this usage is disputed [7] as DCEs (and in some interpretations, BWS [8]) are grounded in random utility theory which makes them conceptually different to conjoint analysis. Therefore, in this paper we use the term attribute-based stated preference (AbSP).

search strategies were peer reviewed by a second information specialist using the Peer Review of Electronic Search Strategies (PRESS) checklist [16, 17].[1]

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. Disagreements were resolved by consensus discussion. This resulted in 16 articles which reported 17 studies.

2.2. Data extraction and analysis

Table 1 lists the studies included for final analysis. Data was extracted using the form in appendix B by one author (EW). Consideration was first given to current practice regarding survey design features (study type, country of origin, participant inclusion criteria, sample size, attribute and levels identification and development) using a comparative summative approach. Detailed consideration was given to attribute development and presentation of information about probabilities, as these are often mentioned as key neglected areas in DCE practice [18].

The quality of each study was scored by one author (EW) by adapting the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) conjoint analysis checklist [19]. The checklist contains 30 items (see Table 2 for a list) which were scored as either 1, if a study reported considering at least some aspect of this item, and 0 if it did not. The final score for a study was then the sum of its scores on each item.

The ISPOR checklist was adapted as it is the only such list specifically on the topic of AbSP studies and was compiled by a committee of experts in the field. It also allowed this study to

[1] Please see Appendix A for full search strategies and date range.

have a secondary aim of testing whether the checklist can serve as an appropriate tool to assess the quality of AbSP studies for future reviews.

Conceptual analysis 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. [20]) 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.

Another potential reason for the preponderance of DMT related studies is that 15/17 (88.2%) were funded by pharmaceutical companies, who have an interest in knowing the preferences of PwMS for different features of DMTs to aid marketing and development of future treatments.

Error! Reference source not found.a 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

Error! Reference source not found.b shows the number of studies using different AbSP methods. The majority (nine, 52.9%) were DCEs.

3.2. Survey population

Error! Reference source not found.c shows the geographical distribution of studies. The most common country in which studies were carried out was the US (seven, 41.2%) with a

further nine (52.9%) spread across Europe (Germany, Italy, the Netherlands Spain and the UK) and a single study in Canada.

3.3. Diagnoses

Error! Reference source not found.d illustrates what diagnoses were used as inclusion criteria. Seven studies (41.2%) included anyone with a diagnosis of MS, and seven (41.2%) included only those with a diagnosis of RRMS. Of the nine (52.9%) studies which clearly indicate that they excluded one or more diagnoses, all excluded PPMS.

3.4. Development of attributes and levels

Error! Reference source not found.a illustrates the sources consulted by studies in developing attributes and levels. Most drew on existing literature in medicine and the social sciences (14, 82.4%) and/or medical professionals (12, 70.6%). Few used qualitative methods to elicit views of PwMS, with only two studies (11.8%) (Lynd, et al. [21], Kremer, et al. [22]) employing focus groups. Seven studies (41.2%) used interviews at some point in the design stage, but this was often only to verify and 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, and the average number for the remaining five is 10.3. Two studies (11.8%) (Wicks, et al. [23] studies 1 and 2) did not state how attributes and levels were developed.

3.5. Survey design

Error! Reference source not found.b shows the number of attributes used by each study. The median number of attributes included in studies is six, with the minimum number used being three and the maximum being 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. [23] studies 1 and 2) do not state what sort of design was used. However, given that a full factorial design would have required 31104 and 768 choice tasks respectively, it is presumed that they also used a fractional factorial design. Five studies (39.4%) selected their

designs based on efficiency, and two (11.8%) explicitly reported using the criterion of D-efficiency.[1] One study (Utz, et al. [25]) 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 what criteria they used to construct their design. **Error! Reference source not found.**a shows the tools used to construct study design. The statistics program SAS (SAS Institute) was the most popular tool 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 to burden participants with [24], and there was considerable variation in the survey length, as can be seen in **Error! Reference source not found.**c which illustrates the number of choice tasks per subject in each study. The median number was 12, with a standard deviation of 14.1. Utz, et al. [25] presented as many as 64 choice tasks, although 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. **Error! Reference source not found.**d shows the number of survey versions used in each study. Two studies (11.8%) (Wicks, et al. [23] 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.

Error! Reference source not found.b 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. [23] study 1) reported power calculations and only six (35.3%) reported other power 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. [26] Bottomley, et al. [27]) reported, in line

[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 [24].

with item 7.2 of the ISPOR conjoint analysis checklist [19], a justification of the chosen mode of administration.

3.6. Attributes

The attributes used by each study were collated and placed in 14 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). The number of studies which include at least one attribute in a given category is displayed in **Error! Reference source not found.**c.

Among the most common attributes were effect on relapse (13, 76.5%), effect on progression (12, 70.6%), as well as severe side effects (also referred to as serious adverse events) (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 four (23.4%) include some other miscellaneous aspect of administration. Six studies (35.3%) looked at attributes related to the alleviation of MS symptoms, but only three (17.6%) included attributes explicitly related to quality of life. Four (23.5%) included attributes related to data from MRI scans. Two (11.8%) include an attribute relating to reproduction (for both men and women).

Eight studies (47.1%) looked at what mode of administration of DMTs PwMS prefer. All included the options of oral medication and injection, though only three out of eight (37.5%) distinguished between subcutaneous and intramuscular injection, and five out of eight (62.5%) included intravenous infusion.[1] All but one of these studies (Utz, et al. [25]) combined mode and frequency of administration into a single attribute with the disadvantage of making it impossible to fully disentangle the effects and gain deeper insight into PwMS' choices, as often mode and frequency of administration are not directly related when making DMT decisions. On the other hand, it has the advantage of "freeing up" an attribute to describe some other aspect of treatment and of a priori ruling out unrealistic combinations such as daily intravenous infusions.

[1] 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.

3.7. Probability

Both the benefits and risks of DMTs are probabilistic in nature [10]. Despite this, 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. **Error! Reference source not found.**d illustrates the methods studies used to represent probability. Only a single study (Johnson, et al. [28]) clearly documented using visual means to convey probabilistic information[1], 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% 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

Figure 5a shows how many studies used a given method of analysis. 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%).

Figure 5b shows how many studies analysed their data using a given software package. Four out of 17 (23.5%) used Sawtooth (Sawtooth Software) and 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

Addressing the needs of individual patients is a crucial part of shared decision making [29], so it is important to go beyond mean preferences to examine the heterogeneity of preferences of PwMS. Only eight studies (47.1%) looked at respondent heterogeneity in some way. The aspects of heterogeneity considered by each study were categorised by one author (EW). **Error! Reference source not found.**c illustrates how many studies examined a given category. Seven out of eight studies (87.5%) tested for the influence of past or current

[1]Wilson, et al. [26] stated that "the visual risk scale was given for reference" but does not elaborate further.

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 or pain and fatigue).

3.10. ISPOR conjoint analysis checklist quality assessment

In general the studies scored well against the ISPOR conjoint analysis checklist, with a median score of 23 out of a possible 30, a minimum of 18 and a maximum of 27.[1] Table 2 gives the mean score achieved by studies on each item of the checklist. The category with the lowest scores was category 3, 'Construction of tasks', with a mean over all three items of 0.18.

Other individual items with a mean score below 0.5 were item 5.3 (mean 0), item 7.1 (mean 0.41) and item 7.2 (mean 0.12). Item 5.3 asks whether a study includes, or reports considering including, a qualifying question to choice tasks indicating strength of preference, confidence in responses, etc. Items 7.1 and 7.2 relate to the previously mentioned issues of sample size considerations and justification of the mode of administration respectively.

4. Discussion

We performed a systematic review of 17 AbSP studies in the field of MS. All but one investigated the preferences of PwMS for aspects of DMTs. This is due to the trade-offs involved in choices between many different treatment options being an ideal topic for AbSP techniques. The vast majority of studies (88.2%) were funded by pharmaceutical companies. This raises the possibility that the aims of this body of literature are biased towards the specific aims of such companies, such as marketing and regulatory approval. The addition of more studies with broader aims, a greater patient focus, and wider applicability would thus be welcome.

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 (for example Cheung, et

[1] An important caveat is that the use of the checklist to judge quality was not overall considered a success, see section 4.

al. [15] found only 62 BWS studies in total published up until April 2016, whereas Clark, et al. [14] found 179 DCEs between 2009 and 2012 alone). It also reflects that the structure of DCEs, 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 forms, for which so far only one drug, ocrelizumab (Ocrevus®) has been approved (for PPMS in the US) [30, 31]. It should be noted that this is still an improvement on the situation a few years ago, and so while focusing exclusively on PwRRMS was appropriate in the past, the anticipated arrival of DMTs for progressive forms of MS means future studies should consider carefully which diagnoses to use as inclusion criteria.

If AbSP studies are properly designed to reflect patient preferences, it is important that the attributes they consider reflect the important issues for PwMS and do not omit vital aspects of the decision making. 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.

If studies are to accurately reflect the range of patient experiences and opinions, they should also examine response heterogeneity. Many of the studies considered here fail to do so, meaning the influence on decision making of factors such as attitude to risk, cognition and previous disease experience are not explored.

Future studies should also consider offering an opt-out option, which only a single study did. Whilst an opt out is not necessarily appropriate for every study, it should be justified given that in situations such as MS some patients choose not to be treated.

The design of a stated preference survey is crucial for the interpretation of its results [19, 32]. 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 even 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 failed to do.

Studies employ a wide range of sample sizes, and such heterogeneity brings into question whether they have recruited appropriate numbers of participants. However, it is difficult to assess whether in general this is or is not the case. Several “rules of thumb” for AbSP sample size exist [33, 34] as well as guides for calculation [35]. Thus power considerations are possible and usually necessary, and should be both undertaken and reported in future studies. If it is not known whether a study is appropriately powered, it causes problems for assessing the quality and validity of its results.

Only two studies reported the reasons for using the mode of survey administration they did, 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, it would be an improvement for future studies to document that concerns over web accessibility, or the unrepresentativeness of a convenience sample from a clinic, for example, have been considered when choosing the mode of administration.

The most common attributes were related to prevention of relapses and progression, and side effects, both mild and severe. These aspects of treatments are probabilistic in nature, yet the majority of studies presented the outcome of treatment decisions as certainties. The appropriate and regular inclusion of probabilistic aspects of DMTs is thus a feature of the literature in need of improvement. Peoples' preferences for probabilistic outcomes are extremely heterogeneous and can have a large 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” [36]) it calls into question the external validity of the results for preferences over real DMTs, whose effects are far from certain.

There is evidence that different ways of presenting probabilities influence individuals' understanding of them [37, 38], and also that understanding can be improved by using graphic presentation of risk [39, 40]. However, those studies which did use probabilistic attributes did not report considering whether their mode of presentation was appropriate, and only a single study conveyed probabilistic information visually. No study looked at 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" [41]), 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 [42, 43]. The impact of risk and uncertainty on PwMSs' decision making regarding DMTs and the optimal way of communicating probabilities thus represent an opportunity for future research.[1] This is particularly the case as studies have revealed cognitive impairments in PwMS [45-49].

Only two studies included an attribute related to reproduction, and in neither of these two did it play a significant role in the analysis. This is despite the high incidence of female PwMS of child-bearing age and that there are also DMTs which are contraindicated for male PwMS trying to conceive [35, 36]. This is of particular relevance given the variation in advice regarding conception, pregnancy and breastfeeding and the lack of clinical research into the influence of DMTs on reproduction [37]. This hence represents an understudied area.

The methods of analysis used by most studies were good, with the majority using sophisticated 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 [50].

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.

We highlight areas in which current practice can be improved. Many studies have not employed qualitative methods to develop their attributes, relying on past literature and clinical opinion. Utilizing qualitative methods to include the views of PwMS is best practice for AbSP and would make the attributes included more relevant to patients' concerns [51]. Various aspects of survey design and development are either neglected or poorly reported.

[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 [44].

Specifically these are power calculations, selecting an appropriate number of attributes per option, number of options per task and mode of survey administration. Studies also rarely report considering an opt out option, despite the ability of PwMS to opt out in real life.

This review highlights gaps in the current literature, particularly the impact of DMT's effects on reproduction on patient preferences. Many studies fail to incorporate the probabilistic nature of DMT's effects, and to date none have examined how different methods of risk communication affect the decision making of PwMS. Likewise, even though many effects of DMTs are uncertain rather than risky, no study so far has incorporated uncertainty.

Our study has several limitations. We have not quantitatively combined the results of studies. We took this decision partly because of a desired 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.

The quality of studies was judged using the ISPOR conjoint analysis checklist [19], which was not wholly successful and is not recommended for future reviews of AbSP studies. Overall there was low variation in study scores as well as low variation in most scores of individual items. The checklist often does 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 qualitative research, systematic literature reviews, interviews and pilot studies. That the checklist is 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.

Due to the focus on details of study design, unpublished studies, mainly from conference proceedings, were excluded, as it was felt that they included insufficient information about the way they were conducted. However, this raises the possibility that the review gives an incomplete picture of current practice.

5. Conclusion

Shared decision making and including patient preference views on treatment are increasingly a part of medicine, including in MS [52, 53]. Thus, it is vital to investigate patient preferences, especially when the experiences of PwMS are highly heterogeneous and there are many treatment options available.

Attribute-based stated preference studies such as discrete choice experiments are increasingly used to measure PwMS' preferences, and can give vital insight into this field. However, several areas of current practice have been highlighted which could be improved, most prominently greater use of qualitative methods in attribute and level development. In addition, reproduction and the presentation of risk have been highlighted as neglected areas that could benefit from future research.

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] 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.

[13] 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.

[14] 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.

[15] 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.

[16] 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.

[17] 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.

[18] 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.

[19] F. Reed Johnson, E. Lancsar, D. Marshall, V. Kilambi, A. Mühlbacher, 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 in Health,* vol. 16, pp. 3-13, 2013.

[20] 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.

[21] 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.

[22] 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.

[23] 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.

[24] 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.

[25] 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.

[26] 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.

[27] 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.

[28] F. R. 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.

[29] 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.

[30] 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.

[31] 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.

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

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

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

[35] 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.

[36] 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.

[37] 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.

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

[39] 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.

[40] 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.

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

[42] 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.

[43] 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.

[44] 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.

[45] 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.

[46] 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.

[47] 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.

[48] 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.

[49] 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.

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

[51] J. Coast, H. Al-Janabi, E. J. Sutton, S. A. Horrocks, A. J. Vosper, D. R. Swancutt, *et al.*, "Using qualitative methods for attribute development for discrete choice

experiments: issues and recommendations," *Health economics*, vol. 21, pp. 730-741, 2012.

[52] 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.

[53] 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.

Figure 1: PRISMA flow diagram

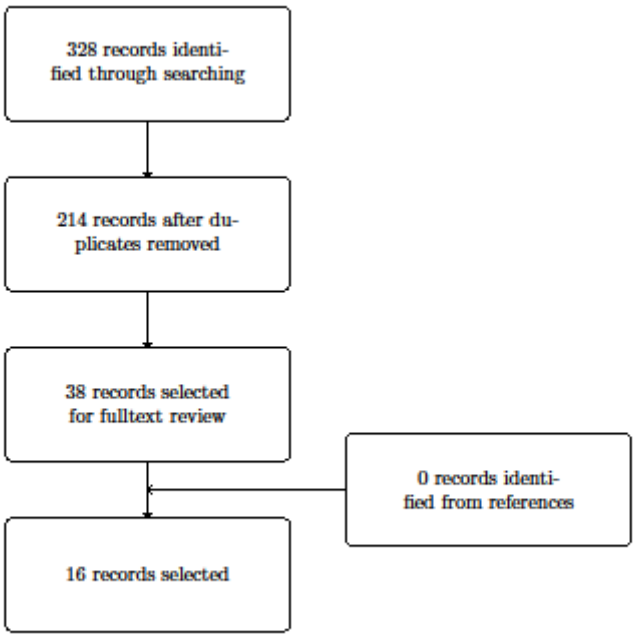

Figure 2: Number of publications of MS attribute-based stated preference studies per year (a), number of studies using a given conjoint analysis method (b), countries in which studies were conducted (c) and number of studies with given MS diagnoses as inclusion criteria (d). DCE = discrete choice experiment, BWS = best-worst scaling, MDU= Multidimensional unfolding, RRMS = relapsing remitting Multiple Sclerosis, CIS = clinically isolated syndrome, SPMS = secondary progressive Multiple Sclerosis.

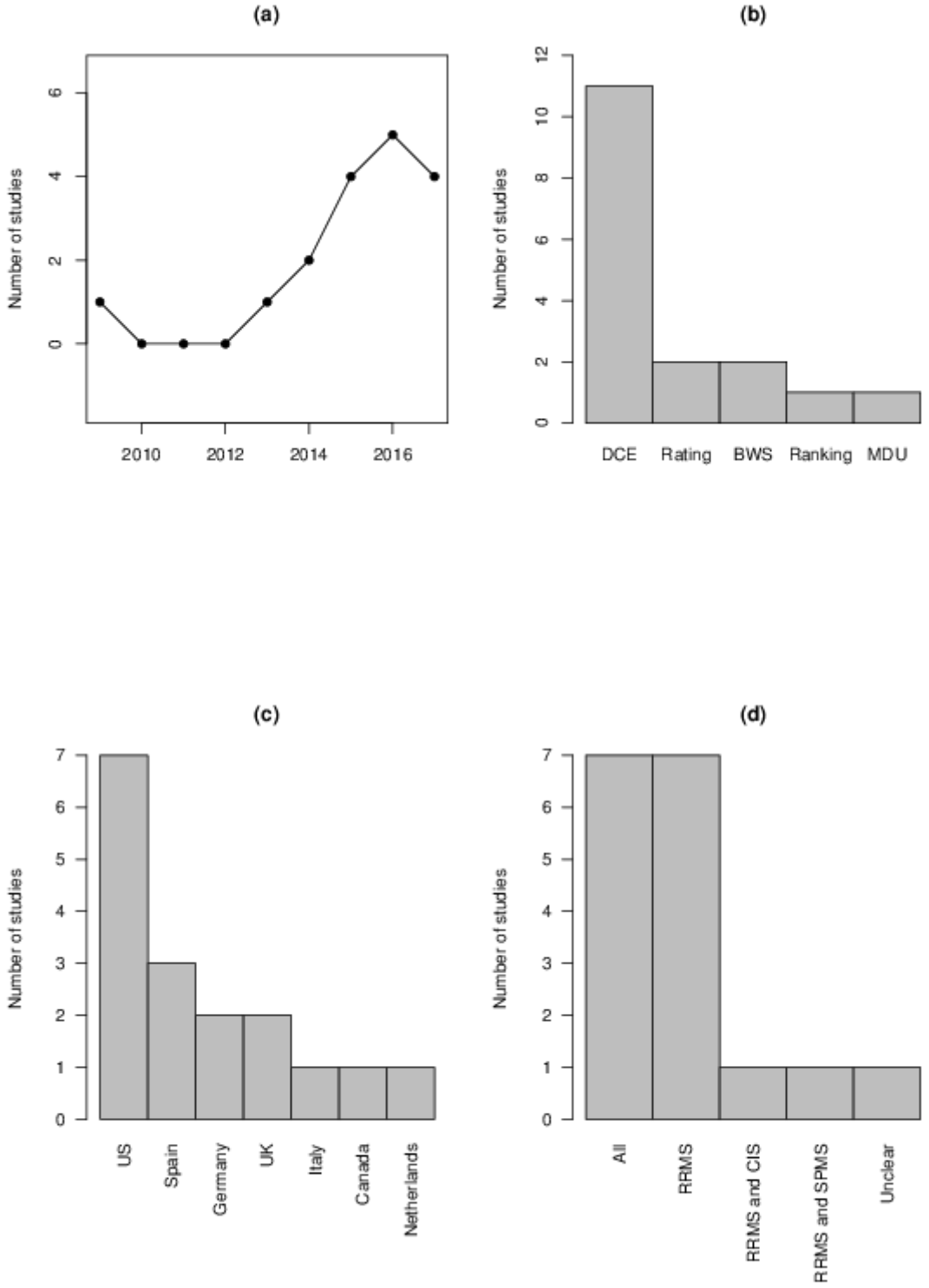

Figure 3: Methods used to develop attributes and levels (a), number of attributes included in each study (b), number of studies including at least one attribute in a given category (c) and number of studies presenting probabilities using a given method (d).

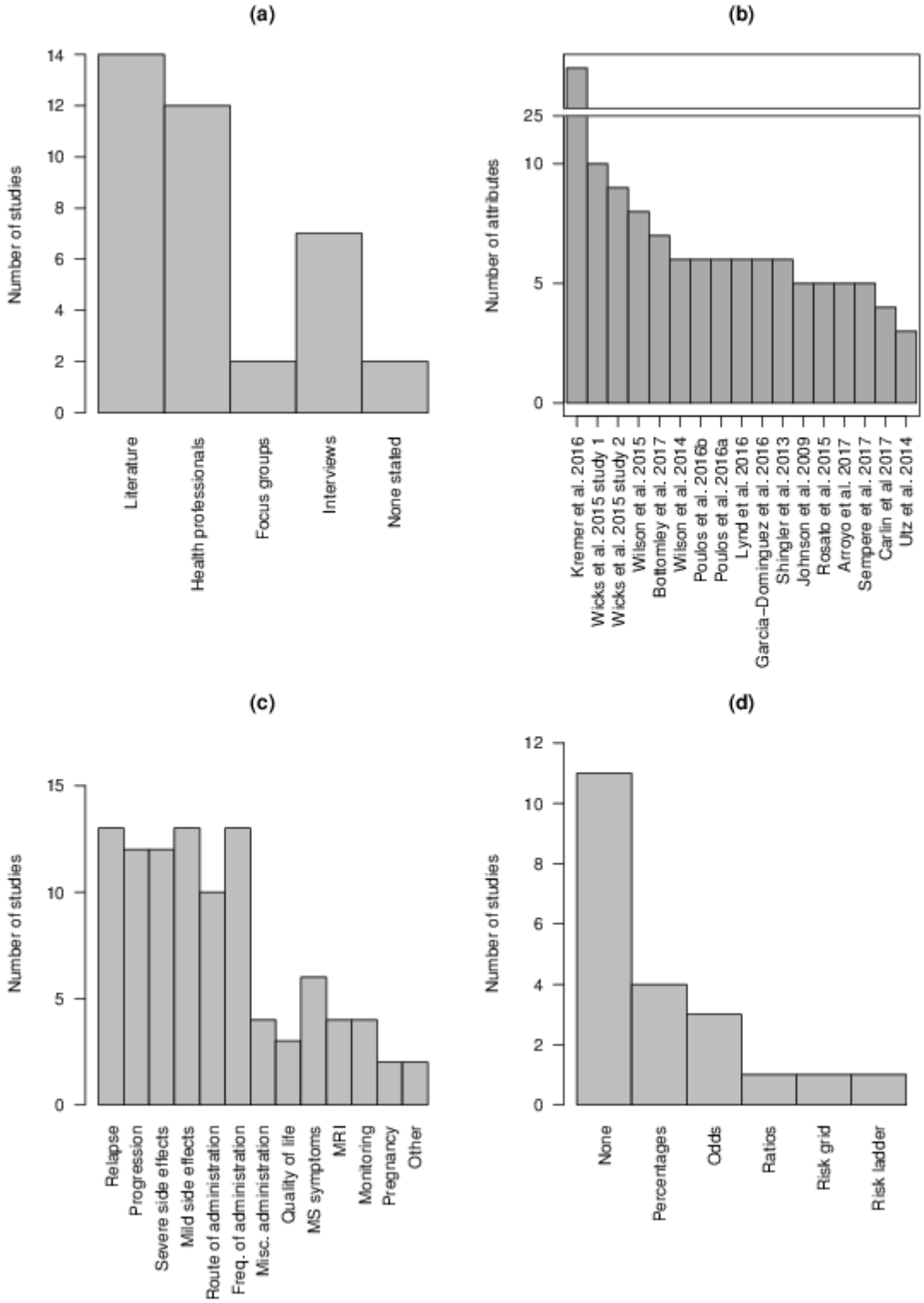

Figure 4: Methods used to construct designs (a), sample sizes for final analysis for each study (b), number of questions answered per subject in each study (c) and number of survey numbers for each study (d).

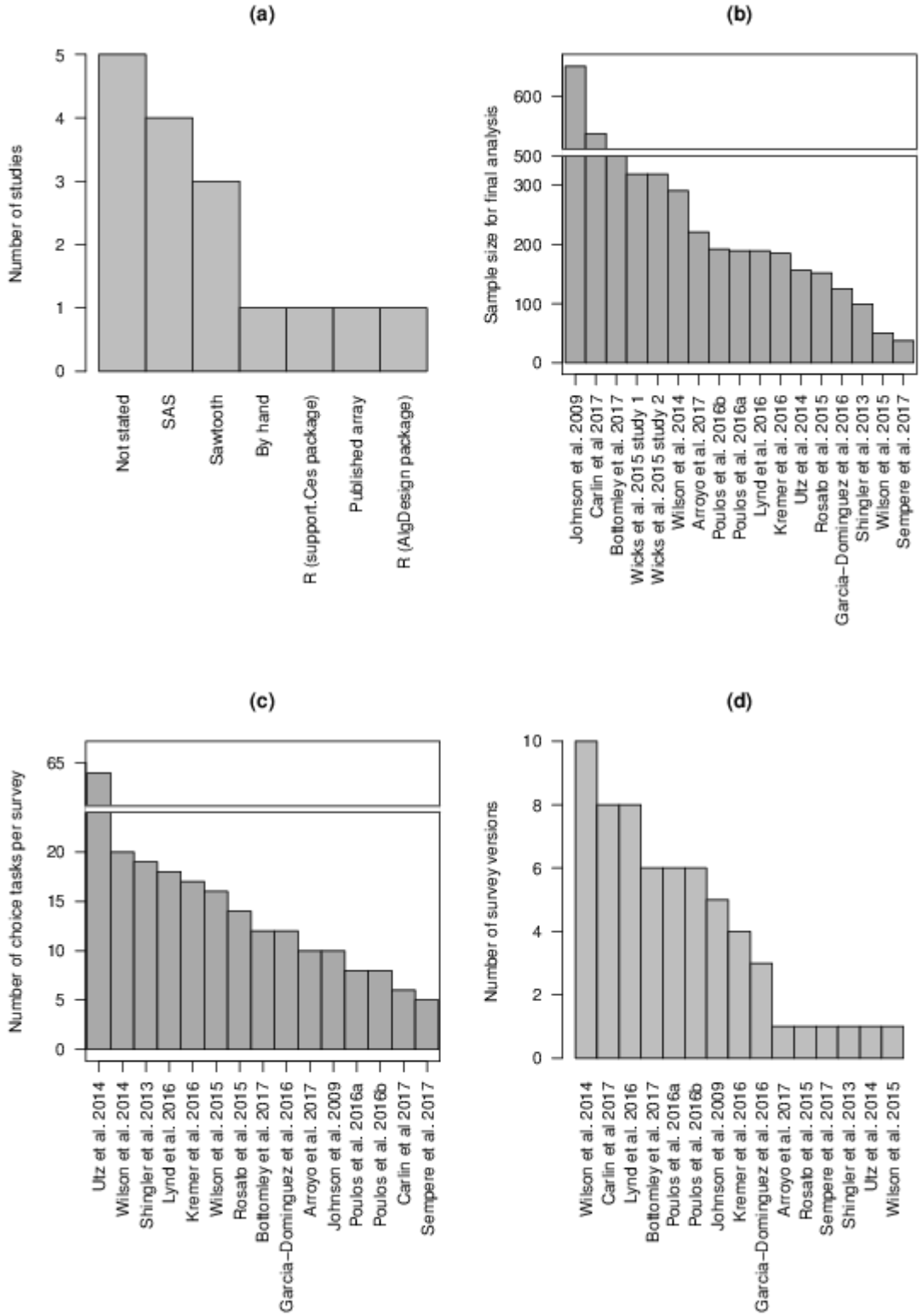

Figure 5: Number of studies using a given analysis method (a), number of studies using a given software package (b) and number of studies examining a given aspect of preference heterogeneity (c).

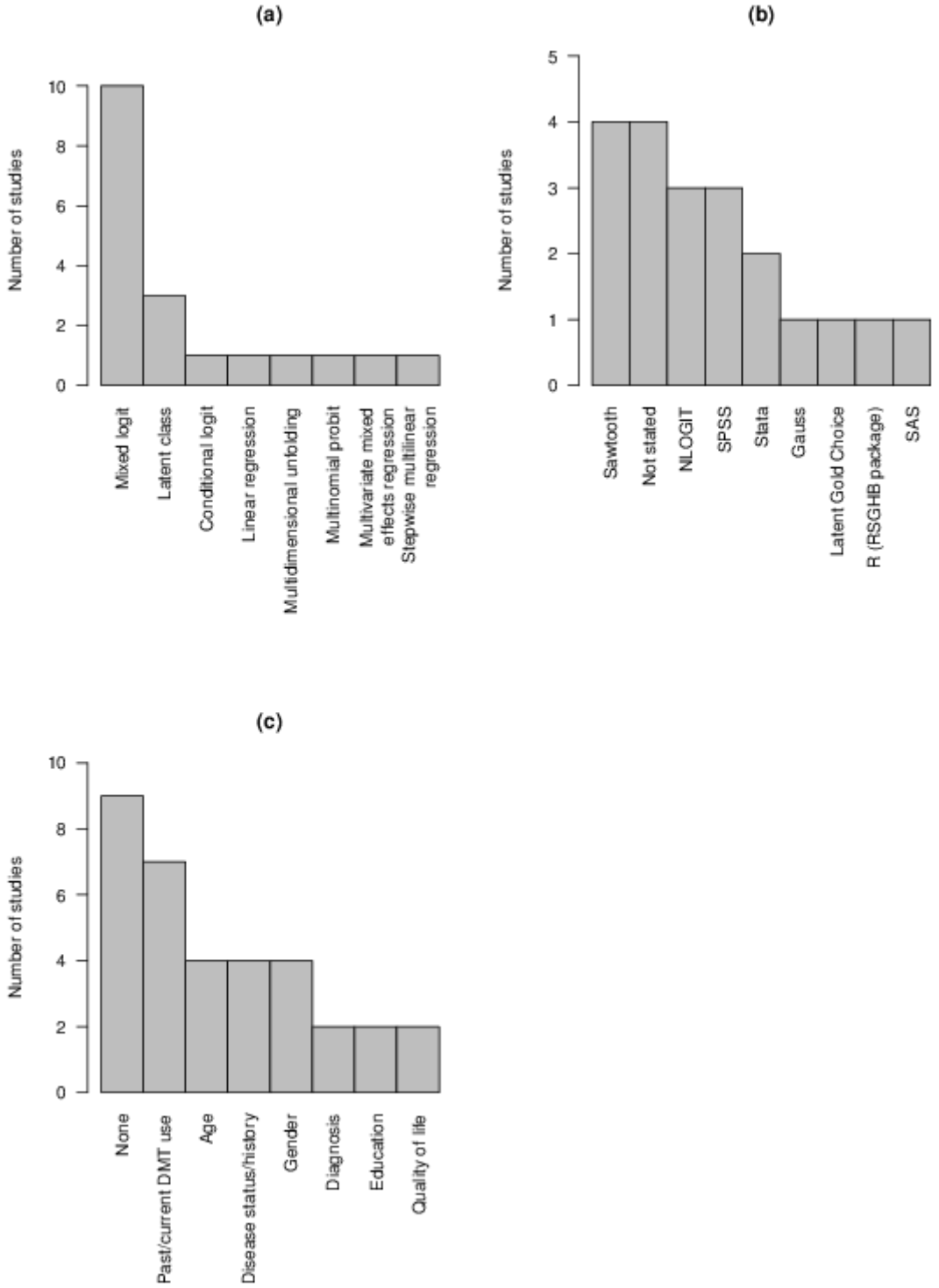

Table 1: Studies included for review. DMT = disease modifying treatment, DCE = discrete choice experiment, BWS = best-worst scaling, MDU = multidimensional unfolding.

Study	Topic	Type	Sample Size	No. of attributes	Mean levels per attribute	Questions per survey
Arroyo et al. (2017)	DMTs	Rating	221	5	2.6	10
Bottomley et al. (2017)	DMTs	DCE	350	7	3.29	12
Carlin et al. (2017)	DMTs	DCE	537	4	5	6
Garcia-Dominguez et al. (2016)	DMTs	DCE	125	6	2.67	12
Johnson et al. (2009)	DMTs	DCE	651	5	4	10
Kremer et al. (2016)	DMTs	BWS	185	27	1	17
Lynd et al. (2016)	DMTs	BWS	189	6	3.17	18
Poulos et al. (2016a)	DMTs	DCE	192	6	3.17	8
Poulos et al. (2016b)	DMTs	DCE	189	6	3.17	8
Rosato et al. (2015)	Quality of life	DCE	152	5	3	14
Sempere et al. (2017)	DMTs	MDU	37	5	1	5
Shingler et al. (2013)	DMTs	DCE	99	6	3	19
Utz et al. (2014)	DMTs	DCE	156	3	2.67	64
Wicks et al. (2015) study 1	DMTs	DCE	319	10	2.9	Not stated
Wicks et al. (2015) study 2	DMTs	Rating	319	9	2.11	Not stated
Wilson et al. (2014)	DMTs	DCE	291	6	3.33	20
Wilson et al. (2015)	DMTs	Ranking	50	8	4	16

Table 2: Mean score for each item of the ISPOR conjoint analysis checklist score

	Checklist item	Mean	(s.e.)
1.1	Were a well-defined research question and a testable hypothesis articulated?	1	0
1.2	Was the study perspective described, and was the study placed in a particular decision-making or policy context?	1	0
1.3	What is the rationale for using conjoint analysis to answer the research question?	1	0
2.1	Was attribute identification supported by evidence (literature reviews, focus groups, or other scientific methods)?	0.88	0.08
2.2	Was attribute selection justified and consistent with theory?	0.76	0.11
2.3	Was level selection for each attribute justified by the evidence and consistent with the study perspective and hypothesis?	0.59	0.12
3.1	Was the number of attributes in each conjoint task justified (that is, full or partial profile)?	0.29	0.11
3.2	Was the number of profiles in each conjoint task justified?	0.12	0.08
3.3	Was (should) an opt-out or a status-quo alternative (be) included?	0.12	0.08
4.1	Was the choice of experimental design justified? Were alternative experimental designs considered?	0.59	0.12
4.2	Were the properties of the experimental design evaluated?	0.59	0.12
4.3	Was the number of conjoint tasks included in the data-collection instrument appropriate?	1	0
5.1	Was there sufficient motivation and explanation of conjoint tasks?	1	0
5.2	Was an appropriate elicitation format (that is, rating, ranking, or choice) used? Did (should) the elicitation format allow for indifference?	1	0
5.3	In addition to preference elicitation, did the conjoint tasks include other qualifying questions (for example, strength of preference, confidence in response, and other methods)?	0	0
6.1	Was appropriate respondent information collected (such as sociodemographic, attitudinal, health history or status, and treatment experience)?	1	0
6.2	Were the attributes and levels defined, and was any contextual information provided?	1	0
6.3	Was the level of burden of the data-collection instrument appropriate? Were respondents encouraged and motivated?	1	0
7.1	Was the sampling strategy justified (for example, sample size, stratification, and recruitment)?	0.41	0.12
7.2	Was the mode of administration justified and appropriate (for example, face-to-face, pen-and-paper, web-based)?	0.12	0.08
7.3	Were ethical considerations addressed (for example, recruitment, information and/or consent, compensation)?	1	0

8.1	Were respondent characteristics examined and tested?	1	0
8.2	Was the quality of the responses examined (for example, rationality, validity, reliability)?	0.53	0.12
8.3	Was model estimation conducted appropriately? Were issues of clustering and subgroups handled appropriately?	1	0
9.1	Did study results reflect testable hypotheses and account for statistical uncertainty?	0.94	0.06
9.2	Were study conclusions supported by the evidence and compared with existing findings in the literature?	1	0
9.3	Were study limitations and generalizability adequately discussed?	1	0
10.1	Was study importance and research context adequately motivated?	1	0
10.2	Were the study data-collection instrument and methods described?	0.88	0.08
10.3	Were the study implications clearly stated and understandable to a wide audience?	1	0

A Search strategy

COCHRANE

Search Name: CRIMSON Dicrete choice experiments
Date Run: 10/07/17 21:56:12.926
Description: 08-02-17

- Cochrane Database of Systematic Reviews : Issue 7 of 12, July 2017 (n = 2)
- Cochrane **Central Register of Controlled Trials : Issue 6 of 12, June 2017** (n = 47)
- Cochrane Methodology Register : 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>

Search Strategy: EMB CRIMSON discrete choice expt MS

---

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 Present> 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>

Search Strategy: PSY CRIMSON discrete choice expt MS

---

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 [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")

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