Health Policy 123 (2019) 118–129

Contents lists available at ScienceDirect

Health Policy

journal homepage: www.elsevier.com/locate/healthpol

Value assessment of disease-modifying therapies for Relapsing-Remitting Multiple Sclerosis: HTA evidence from seven OECD countries

Erica Visintin*, Michela Tinelli, Panos Kanavos

Department of Health Policy, Medical Technology Research Group, LSE Health, London School of Economics, Houghton Street, London WC2A 2AE, England, United Kingdom

ARTICLE INFO

Article history:
Received 7 February 2018
Received in revised form 24 June 2018
Accepted 30 August 2018

Keywords:
Health Technology Assessment
Multiple sclerosis
Disease modifying treatment
Pharmaceutical policy
Scientific evidence
Social value judgments
Evidence uncertainty

ABSTRACT

This study systematically compares HTA recommendations on a number of disease–modifying therapies for patients with Relapsing-Remitting Multiple Sclerosis. We analysed publicly available HTA reports for nine medicine-indication pairs across seven OECD countries using a methodological framework enabling systematic analysis of HTA recommendations. The analysis was conducted based on a number of value dimensions, including clinical and economic variables, as well as several other dimensions of value beyond cost-effectiveness. The material was qualitatively and quantitatively coded following the different stages of HTA decision-making process. Fifty-seven medicine-indication pairs were assessed across the study countries. Of those, eight medicine indication-pairs reported diverging HTA recommendations. Although HTA recommendations were based on the same evidence submitted in most cases, significant variations were identified in interpretation and acceptance of evidence resulting in different uncertainties raised and different ways of addressing them. Uncertainties arose both in terms of the clinical and the economic evidence, including the design of key trials or the data quality in economic models. Beyond costs and effects, additional dimensions of value had an impact in the direction of recommendations, however with different magnitude across countries. We show that there is heterogeneity across countries in HTA for evaluating DMTs for RRMS with a lack of standardised methods in evaluating clinical and economic evidence and the use of social value judgments to inform decision-making.

1. Background and objectives

Multiple Sclerosis (MS) is a fairly commonneurodegenerative disease affecting the central nervous system (CNS) in young adults and causing permanent disability [1]. MS affects nearly 2.3 million people, has a prevalence of 30/100,000 [2] globally, with prevalence being higher in countries further from the equator such as Canada (291/100,000), the UK (203/100,000) and Germany (128/100,000) [2]. MS is also associated with a high direct and indirect cost of illness, depending on the severity of the disease [3]. Medicines are the main cost driver for MS patients with low disease severity, whereas the costs relating to productivity losses are associated with a significant impact on families and society and are the main cost drivers in later stages of the disease [3].

The past two decades have witnessed remarkable advances in treatment options for MS with currently eleven disease modifying therapies (DMTs) being available that have convincingly altered the short- and medium-term natural history of the disease [4], reducing the rates of relapse in patients with relapsing-remitting MS (RRMS), the associated permanent neurological damages and slowing MS progression [5–10] particularly when treatment is initiated at early stage [11–13]. DMTs can be costly, however, considering the increased costs associated with relapse occurrence and increasing disease severity, DMTs have been used early in an attempt to delay disease progression, ultimately contributing to long-term medical and societal benefits, including economic savings [14]. The early accessibility to a wide range of DMTs might provide both alternative options for MS care management to fit patient needs and secure positive outcomes for patients and society.

Budget constraints across health systems question the overall financial sustainability and equality in access to care. A fundamental objective of Health Technology Assessment (HTA) is to address these concerns and support decision-makers with choices regard-

* Corresponding author.
E-mail addresses: e.visintin@lse.ac.uk (E. Visintin), m.tinelli@lse.ac.uk (M. Tinelli), p.g.kanavos@lse.ac.uk (P. Kanavos).

https://doi.org/10.1016/j.healthpol.2018.08.019
0168-8510/

ing the services and products that are made routinely available by national healthcare systems. In the context of MS, there is a need to understand how evaluation of DMTs is conducted and which HTA decision-making factors may result in differential outcomes impacting access to these treatments across countries.

This paper systematically compares HTA recommendations on a sample of DMTs across different countries, in order to understand the criteria driving value assessments, what might explain any differences in HTA recommendations and what the policy implications of such differences might be. Although several studies have investigated the impact of HTA processes across different countries and diagnoses [15–20], no study has analysed HTA outcomes on MS DMTs. The paper provides a systematic analysis of similarities and differences in HTA recommendations across settings, the reasons for these, their implications for coverage and draws broader conclusions about policy directions based on the results of this analysis. The results enable general inferences to be drawn about the evidence considered when evaluating MS DMTs, and highlight the key elements that need to be included when appraising chronic conditions.

2. Methods

2.1. Analytical framework

The collection and analysis of HTA data were based on a standardized analytical framework developed and applied to compare and evaluate HTA evidence across settings using a mixed methods approach [15]. Based on that, the four stages selected to perform the analysis matched the key components of HTA notably: (1) the type of clinical and economic evidence considered across therapeutic options and countries; (2) the interpretation of the clinical and economic evidence; (3) the inclusion of social value judgements and stakeholder opinion and how they relate to the first two components throughout the assessments; and (4) the influence of all these elements on the appraisal and the final recommendation. This multi-stage methodology contributes to identifying: (a) the key factors influencing HTA decision-making processes, (b) how these factors interact with each other (e.g. stakeholder input and concerns raised), and (c) how these factors relate to a certain outcome, what determines similarities and differences in appraisals and, ultimately, coverage recommendations.

2.2. Data sources and inclusion criteria

The data used for this purpose were extracted from publicly available HTA reports available through national HTA agency websites. The MS medicines considered in this analysis were DMTs used in the treatment of Relapsing-Remitting-Multiple-Sclerosis (RRMS) the most common sub-type of the MS disease affecting about 85% of people with MS. This also included cases of rapidly evolving severe (RES) RRMS affecting 15–20% of people with RRMS [2]. RRMS is characterized by clearly defined attacks of new or increasing neurologic symptoms and can be categorized as either active (with relapses and/or evidence of new MRI activity) or not active, as well as worsening (a confirmed increase in disability over a specified period following a relapse) or not worsening. RES-RRMS is characterised by two or more disabling relapses in one year and evidence of increasing lesions on two consecutive MRI scans.

The study countries (England, Scotland, Sweden, France, Germany, Canada, Australia) were selected based on a number of criteria, notably whether: (1) they had well-established HTA agencies and processes; (2) they deployed different HTA approaches (clinical and cost-effectiveness analysis or comparative clinical benefit assessment); (3) they represented a sample of the different approaches to HTA (e.g. health service, societal); (4) their HTA reports were publicly available.

Across all study countries, we identified a single authority from which we retrieved the HTA reports, with the exception of Germany, where two authorities (one advisory and one regulatory) were considered. Eight HTA bodies were selected: The Canadian Agency for Drugs and Technologies in Health (CADTH, Canada); Haute Autorité de Santé (HAS, France); Institut fuer Qualitaet und Wirtschaftlichkeit im Gesundheitswesen (IQWiG, Germany); Gemeinsamer Bundesausschuss (G-BA, Germany); National Institute for Health and Care Excellence (NICE, England); Scottish Medicines Consortium (SMC, Scotland); Pharmaceutical Benefits Advisory Committee (PBAC, Australia); and the Dental and Pharmaceutical Benefits Agency (TLV, Sweden). In Canada, the Canadian Agency for Drugs and Technologies in Health (CADTH) was selected as being representative of the Canadian federal system for producing recommendations using comparable criteria and methods as other agencies in the sample. In the case of Germany we considered both IQWiG and G-BA; the latter is the critical HTA body (with a regulatory capability), while the former has advisory capability and makes a recommendation on the basis of a systematic and structured report for non-orphan drugs. Across settings the data retrieved related to the latest or final HTA report for a given medicine-indication pair and did not include (but accounted for) likely earlier rejections.

The unit of analysis of the study was the medicine-indication pair, comprising all the different types of RRMS, including RES-RRMS. The inclusion criteria for selecting the different medicine-indication pairs were: (1) the medicine was appraised by at least four study countries; (2) any recommendation produced following assessment by an HTA agency was published before 31 July 2017; (3) the assessment across different countries should have the same indication (e.g. tecfidera has the same indication for the treatment of active relapsing-remitting multiple sclerosis across all HTAs); (4) the medicines should have a Marketing Authorization (M.A.) across all study countries.

2.3. Data extraction, coding and analysis

First, a database was created to identify and select all suitable medicine-indication pairs considered in our analysis for the treatment of RRMS. In order to enable comparisons and draw inferences about the extent to which commonalities and differences existed among HTAs, a number of parameters were extracted across all HTA reports to account for and compare the different criteria considered at each stage of the decision-making process (Table 1). The extracted parameters were included in a database and comprised the following: (1) the molecule name; (2) the branded medicine name; (3) the HTA recommendation issued in each study country (List (L), List with constraints (LWC), and Do Not List (DNL) or the level of added benefit for Germany and France, ranging from 1 to 6 and 1 to 5 respectively); if a drug-indication pair was not submitted for assessment, the designation Not Assessed (NA) was applied; (4) the M.A. date; (5) a number of variables outlining clinical evidence (e.g. type of clinical evidence, trial design and phase, comparator, type of endpoint); (6) assessment of the clinical evidence by HTA agencies with particular reference to clinical uncertainties, particularly those arising from study design, poor data, the size of the clinical benefit, the relationship of the treatment to current clinical practice and population generalisability; (7) a number of variables outlining economic evidence submitted (e.g. type of economic model and comparator used); (8) assessment of the economic evidence by HTA agencies and, specifically, uncertainties arising from the model and modelling used, the comparator, the cost and utility data used, the clinical evidence used in the economic model, the clinical assumptions made, and the relationship

Table 1
The HTA-related parameters recorded in the analysis of RRMS DMTs across seven OECD countries.

Parameters	Data recorded	
Final recommendation	Listing decision (CADTH, NICE, PBAC, SMC, and TLV) or Level of benefit (HAS, IQWIG and G-BA); decisions classified as: L (List), LWC (List with Criteria), DNL (Do Not List); NA (not assessed)	
Date of recommendation	Date of HTA recommendation	
Trial design	Pivotal/non-pivotal; Type of design (randomized, non-randomized, observational, etc); Phase (I/II/III/IV/V); Blinding (double blind/open-label); Comparator (active, placebo, no comparator); Primary/Secondary endpoints; and HRQoL endpoints	
Economic model	Type of model (CUA, CMA, CBA or CCA); Comparator	
Clinical uncertainties	Study design: comprises all the concerns raised across trial design (blinding, phase and clinical or surrogate endpoints, length of trial and follow up period, sample size, comparator). Poor data: comprises all the concerns raised around the lack of long-term data or data for specific population subgroups. Size of clinical benefit: comprises all the concerns raised around the magnitude of clinical benefit (e.g. is too little or confounded by other factors that are not related to the clinical design) the safety information, and the HRQoL data. Relationship to current clinical practice: The clinical trials included do not reflect the current clinical practice (e.g. administration route or pre- and concomitant medication or a different use of the resource of the health system) Population generalisability: all concerns raised around the generalizability of the population used in the clinical evidence to the country of the HTA body	
Economic uncertainties	Model and modelling: comprises concerns around the use of a certain model (cost-minimization or cost-utility etc.) or the modelling used (Markov/ partitioned survival model), or the extrapolation technique used for the clinical data. Comparator: comprises concerns around the comparator used in the economic analysis. Cost and utility data concerns: concerns around the data used to build the model (e.g. costs and utility values) leading to overestimation or underestimation of the ICER. Clinical evidence used in the economic model: comprises concerns around the suitability of the trial used to populate the clinical input in the economic model, the magnitude of the clinical benefit captured and the population included. Clinical assumptions: concerns around the clinical assumptions used in constructing the economic model (e.g. natural progression of disease time) Relationship to current clinical practice: the assumptions around frequency of dosage, resource used by the NHS and the concomitant medications used.	
Social value judgements	All the elicited (e.g. human dignity and solidarity principles in Sweden; end-of-life criteria in the UK) and not-elicited considerations (e.g. the emotional strain put on patient families, administration advantages, financial impact on patients and their surrounding) made by the HTA body in order to capture the full impact of the disease from a broader societal perspective.	
	Chronic/Incurable/ onset in young age	Refers to the duration of the disease and the age at onset.
	Small population/Rare disease	Refers to cases when the rarity of the disease or the orphan status of the drug were recognised.
	Administration route and frequency	Refers to the route and the frequency of administration of the treatment.
	Unmet need	Refers to whether the unmet need for new treatments was recognised (e.g. few or no alternatives exist, need for additional treatments).
	Social life	Refers to the ability of patients and their families/caregivers to have a normal social life, including leisure activities.
	Financial burden	Financial burden created by the disease to patients or families.
	Emotional burden	Emotional burden created by the disease to patients or families.
	Impact on the organization of care	The wider impact on the health care system (e.g. administration in hospital setting or provision of specific testing)
	Ethical/equality issues	The possible inequalities arising from the reimbursement of the treatment including those associated with gender, age, race, disability and/or socioeconomic status.
	Stigma	The nature of the disease and its impact on patients and their surroundings' quality of life (e.g. anxiety from the disease leading to social stigma).
	Productivity	Any information provided about the indirect benefits from the treatment, (e.g. being able to return to work, improving functional capacities) and also the ability of carers to perform regular work activities.
	Disabling nature	Refers to the disabling nature of the disease as well as the capacity of the treatment to reduce the disability/slow the progression of the disability
Stakeholder input	The extent and effect of the involvement of external stakeholders (e.g. patient representatives, clinicians or experts) on the final recommendation.	

Source: The authors based on adaptation and enhancement of the conceptual framework by Nicod and Kanavos [15].

to current clinical practice; and (9) any social value judgements (SVJs) used to inform decision-making, whether these were elicited (e.g. such as the human dignity and solidarity principle in Sweden), or non-elicited, (e.g. such as emotional strain on families, administration advantage, financial impact on patients and families). The inclusion of SVJs in the decision-making process is particularly significant in the current context. Clinical and economic evidence is often incomplete or of low quality and, therefore, additional factors, such as clinical, social or ethical parameters are considered by HTA agencies as relevant elements when considering the cost effectiveness of new treatments. Such factors are increasingly been included in HTA decision-making across countries [4].

Nvivo 11 was used to code the above variables and transfer the qualitative results into a quantitative Excel file and attribute to each code item a number. Each code item was composed by text segment referring to the relevant concept in the evaluation of a medicine (e.g. A naïve indirect comparison was recorded under the code "indirect comparison") and was incorporated in a macro category code (e.g. all the different types of clinical trials were included under the macro code "clinical evidence"). We looked for the incidence of each concept (e.g. how many times in a report clinical uncertainties around clinical benefit were raised and addressed) across HTA reports and countries and produced the sum of occurrences for each concept. By doing this, we were able to record the frequency of incidence of a specific code or code group as well as study associations across code items.

This approach enabled to translate the qualitative value of the criteria identified into a quantitative system, weigh the relevance

of specific criteria looking at both specific sub-groups (medicines or countries) and the whole sample, and build the final datasets used for the quantitative analysis. Coding was conducted systematically and homogeneously across all medicine-indication pairs, such that codes were also comparable across study countries. The information coded informed three separate analyses: first, a qualitative analysis to understand the HTA process in each study country; second, a quantitative descriptive analysis to identify trends across countries; and, third, the triangulation of the qualitative and quantitative data complementing the previous two analyses and allowing to identify how individual HTA criteria could influence the final recommendation, measure the extent of differences across cases (in terms of DMTs or countries) and how these contributed to explaining different recommendation outcomes.

3. Results

3.1. RRMS Medicines and HTA recommendations

A total of 11 DMT medicines were identified, although two of them, i.e. IFNβ 1a IM (Avonex) and IFNβ 1a SC (Rebif), were excluded from final analysis as they were appraised by three HTA agencies only. From the sample of nine DMTs considered in this study, natalizumab (Tysabri) was the only treatment indicated for RES-RRMS (see Table 2).

In total, 56 HTA reports were identified relating to a total of 57 medicine-indication pairs across the study countries (n = 9 for each of HAS and PBAC; n = 8 for each SMC, NICE and TLV; n = 7 for CADTH; n = 3 for IQWiG; and n = 5 for G-BA); in the case of NICE, there was a multiple technology assessment (MTA) of Glatiramer acetate and IFNβ 1b SC, therefore a single report assessed 2 medicines. A complete list of the reports with their respective web link is shown in Appendix 1 in Supplementary data.

In ten cases (3 each by TLV and PBAC, 2 by CADTH and 1 by SMC and HAS) HTA bodies produced re-assessments based on a re-submission by the manufacturer of new prices or new clinical evidence, which overturned earlier recommendations from DNL to LWC (8 times) or to L (twice). In four cases re-assessments were conducted because the economic case was not made by the manufacturer in earlier submissions (1 by PBAC; 1 by SMC; and 2 by TLV); in three cases re-assessments took place because clinical evidence was insufficient in earlier submissions (2 by PBAC; and 1 by CADTH) rendering a reliable economic evaluation impossible. Finally, in two cases (1 by CADTH and 1 by TLV) the manufacturer withdrew the first submission without specifying the reasons behind the decision while in one case (HAS) the manufacturer submitted new clinical data for the continuation of the treatment in the national reimbursed list.

In 68% of cases (n = 39) the studied medicine-indication pairs received a restricted recommendation (LWC), followed by 25% receiving a listing recommendation (L). Only 7% were rejected (DNL) across the entire sample. Of the 9 DMTs, three showed opposite HTA recommendations between the country settings (varying from DNL, to LWC or L), five showed discordant recommendations, but positive and in the same direction (either LWC or L), one had LWC recommendations only (daclizumab) (Table 2). The three opposite recommendations were related to assessments in different time periods. IFNβ 1b SC was rejected by NICE, PBAC and SMC and accepted by HAS and TLV. HAS, SMC and PBAC assessed the drug in 2007 whereas NICE and TLV in 2002. Teriflunomide was rejected by CADTH, accepted by TLV, and accepted with restrictions by all other agencies, but the time difference between assessments was less than 12 months. Glatiramer acetate received a rejection by NICE, was accepted by PBAC and SMC and was accepted with restrictions by TLV and HAS. Between NICE and the other agencies there was a significant time difference (9.5 years with HAS and 13.5 years with TLV). In 13 out of 30 mildly discordant recommendations (either LWC or L), one or more economic restrictions were applied (e.g. PAS by NICE and SMC or a risk sharing agreement (RSA) by PBAC). Looking at the clinical restrictions applied, 21 were pertaining to a specific clinical subgroup with 19 being the same across agencies and 16 were pertaining to specialist use. The longest time difference across mildly discordant recommendations was seen in the case of alemtuzumab with HAS last assessing the drug in January 2016 and SMC firstly assessing it in April 2014 (20 months).

3.2. Clinical evidence

A total number of n = 217 clinical studies (of which 102 involved a comparator) across nine DMTs and seven countries were identified that supported HTA submissions. Of these, 40% were phase III trials, the remainder being indirect comparisons (33%), observational studies (11%) and phase II trials or trial extensions (8%). HAS considered the highest number of clinical studies (n = 48), followed by NICE (n = 42), PBAC (n = 37), SMC (n = 31) and TLV (n = 20). G-BA and IQWIG considered only eleven and five clinical studies respectively across the 5 and 3 medicine-indication pairs they examined. Fig. 1 shows the distribution of the clinical study types across the study countries.

Of all clinical studies (phase II or phase III), 58% (n = 60) were placebo-controlled whereas 42% (n = 44) reported a direct comparator, such as Beta-interferon therapies or other DMTs therapies as comparator.

When we compared the type of evidence used by HTA agencies for the same medicine-indication pair, we found that across all medicine-indication pairs agencies considered different clinical evidence. This had an impact on the final HTA recommendation in some cases (n = 6 cases out of a total of 56 HTA reports). In the case of natalizumab the different restrictions applied by HTA agencies resulted from different forms of clinical evidence. HAS considered a phase III study (SENTINEL) comparing natalizumab in combination with beta interferon against beta interferon alone, but this evidence was excluded/not considered by NICE and CADTH and was used as supportive evidence only by SMC and PBAC. This was due to the possibility of progressive multifocal leukoencephalopathy reported in the SENTINEL trial. By contrast, HAS included this trial in its evaluation and requested to receive evidence about the conditions of use in concomitant treatments of natalizumab and beta interferon. CADTH simply acknowledged the harms associated with natalizumab use (e.g. progressive multifocal leukoencephalopathy) but recognized the unmet need for therapeutic options for treatment-experienced patients, leading to restrictions for the drug to be available as third line treatment or for intolerant patients.

3.3. Clinical endpoints

Clinical endpoints were grouped into three types, notably: primary (e.g. annual relapse rate); secondary (endpoints analyzed post hoc for which the trial may not have been powered or lacked randomization, e.g. percentage change from baseline in T2-hyperintense lesion volume); and health-related-quality-of-life (HRQOL), (e.g. Multiple Sclerosis Impact Scale [MSIS-29]). Substantial differences were seen in the cumulative number of clinical endpoints (as well as in the type) considered by the study countries. When looking at the overall data reports, HAS included the highest number of endpoints (n = 166 endpoints across the 9 medicine-indication pairs: 66% primary, 30% secondary, 4% HRQOL) followed by NICE (n = 149; 55% primary, 32% secondary, 13% HRQOL), PBAC (n = 80; 61% primary, 39% secondary), SMC (n = 75; 45% primary, 45% secondary, 11% HRQOL), CADTH (n = 58; 50% primary, 28% sec-

Table 2
HTA recommendations for RRMS DMTs across seven OECD countries (January 2002–June 2017).

Molecules name (branded name)	HTA Recommendations and associated dates															
	Australia (PBAC)		Canada (CADTH)		England (NICE)		France (HAS)		Germany (IQWiG)		Germany (G-BA)		Scotland (SMC)		Sweden (TLV)	
	Date	Outcome	Date	Outcome	Date	Outcome	Date	Outcome	Date	Outcome	Date	Outcome	Date	Outcome	Date	Outcome
Alemtuzumab (Lemtrada)	Jul-14	LWC	Nov-14	LWC	May-14	L	Jan-16	ASMR V, 65% reimbursement SMR Important	NA	NA	Sep-16	Added benefit not proven (Level V)	Apr-14	L	NA***	NA***
Daclizumab (Zinbryta)	Nov-16	LWC	Jun-17	LWC	Apr-17	LWC	Jan-17	ASMR V 65% reimbursement SMR Important	NA	NA	NA	NA	Mar-17	LWC	Feb-17	LWC
Dimethyl fumarate (Tecfidera)	Jul-13	L	Sep-13	LWC	Aug-14	LWC	May-14	ASMR V 65% reimbursement SMR Important	Jul-14	Added benefit not proven (Level V)	Oct-14	Added benefit not proven (Level V)	Jul-14	L	Aug-14	LWC
Fingolimod**** (Gylenia)	Mar-11	LWC	Nov-11	LWC	Apr-12	LWC	Jul-11	ASMR IV, 65% reimbursement SMR Important	Nov-12	Hint of a minor added benefit (Level III)	Mar-13	Hint of a minor added benefit (Level III)	Aug-14	LWC	Aug-11	L
Glatiramer acetate (Copaxone)	Mar-15	L	NA	NA	Jan-02	DNL**	Apr-11	ASMR V, 65% reimbursement SMR Important	NA	NA	NA	NA	Nov-15	L*	Jun-15	LWC
IFNβ 1b SC (Betaferon, Extavia)	Mar-07	LWC	NA	NA	Jan-02	DNL**	Jul-07	ASMR III, 65% reimbursement SMR Important	NA	NA	NA	NA	NA	NA	Jul-02	LWC
Natalizumab***** (Tysabri)	Nov-07	LWC	Feb-09	LWC	Aug-07	LWC	Jan-07	ASMR III, SMR Important	NA	NA	Oct-08	Added benefit not proven (Level V)	Aug-07	LWC	Dec-06	L
Peginterferon beta-1a (Plegridy)	Nov-14	LWC	Jun-15	LWC	NA	NA	Apr-15	ASMR IV, 65% reimbursement SMR Important	NA	NA	NA	NA	Dec-12	L	May-15	L
Teriflunomide (Aubagio)	Jul-13	LWC	Jun-14	DNL	Jan-14	LWC	Mar-14	ASMR V, 65% reimbursement SMR Substantial	Jul-13	Added benefit not proven (Level V)	Mar-14	Added benefit not proven (Level V)	Jan-14	LWC	Jun-16	L

Source: The authors from publicly available HTA reports.
CADTH: Canadian Agency for Drugs and Technologies in Health; HAS: Haute Autorité de Santé; IQWIG: Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen; G-BA: Gemeinsamer Bundesausschuss; NICE: National Institute for Health and Care Excellence (NICE); PBAC: Pharmaceutical Benefits Advisory Committee; SMC: Scottish Medicines Consortium; TLV: Dental and Pharmaceutical Benefits Board; L: List; LWC: List with constraints DNL: do not list; RRMS: Relapsing-Remitting Multiple Sclerosis; RES-RRMS: Rapidly evolving severe; * abbreviated decision.**The drug were reimbursed through risk-sharing scheme developed by the Department of Health and they underwent a Multiple technological assessment (MTA); ***TLV published only a Health-economic assessment; **** Fingolimod is indicated for the treatment of highly active RRMS; ***** Natalizumab is indicated for the treatment of Rapidly evolving RRMS; NA: not assessed. Countries are shown in alphabetical order.

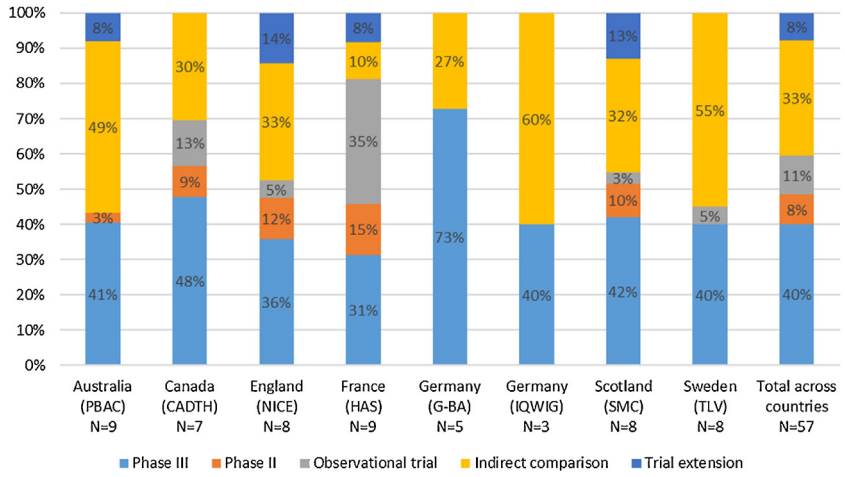

Fig. 1. The types of clinical studies related to RRMS and their contribution to the HTA process across seven OECD countries.
CADTH: Canadian Agency for Drugs and Technologies in Health; HAS: Haute Autorité de Santé; IQWIG: Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen; G-BA: Gemeinsamer Bundesausschuss; NICE: National Institute for Health and Care Excellence (NICE); PBAC: Pharmaceutical Benefits Advisory Committee; SMC: Scottish Medicines Consortium; TLV: Dental and Pharmaceutical Benefits Board; N indicates the number of DMT for RRMS assessed by HTA bodies in each of the study countries. Countries are shown in alphabetical order.
Source: Authors' compilations from clinical evidence presented in HTA reports.

ondary, 22% HRQOL), G-BA (n = 64; 45% primary, 34% secondary, 20% HRQOL), TLV (n = 40; 64% primary, 33% secondary, 3% HRQOL) and IQWIG (n = 19; 58% primary, 11% secondary, 32% HRQOL). The discrepancy in the number of endpoints considered is due to two main factors: first, the difference in the number of clinical studies considered; and, second, the number of subgroups considered by HTA agencies. HAS and NICE assessments reported a greater number of subgroup analyses (17 and 15, respectively), compared with G-BA and SMC reporting respectively only 10 and 6 sub-group analyses, PBAC and CADTH only 5 each, IQWIG only 1 and TLV none; HAS and NICE studied a higher number of secondary endpoints: (33 and 27 respectively), compared with PBAC (22); SMC (18); CADTH (12) and TLV (3). In the case of IQWiG the number of endpoints and sub-group analyses should be considered with caution due to the low number of assessments included in the analysis (only 3 drug-indication pairs) and the fact that IQWIG in their summary do not report all the sub-group analyses required by German regulation and this is more extensively done by the G-BA, which is confirmed by our analysis.

When examining the individual medicine-indication pairs, the greatest differences in endpoints considered across HTA agencies was reported for alemtuzumab, with PBAC considering evidence from 32 primary endpoints, NICE (22), followed by G-BA, HAS, SMC, TLV and CADTH, reporting only twelve, nine, seven, four and two, respectively. This difference was related to the different number of indirect comparisons considered (5 by PBAC; 3 by NICE and G-BA ; 2 by SMC; and 1 by each of HAS and CADTH) leading to an estimation of the co-primary endpoints (onset of sustained accumulation of disability and relapse rate) with multiple analysis and sub-group analysis. Importantly, however, the variation in the number of endpoints considered did not lead to differences in recommendations across HTA agencies, but was associated with the time difference in assessing alemtuzumab across HTA agencies as well as the direction in the recommendation, including restrictions made.

3.4. Uncertainties relating to clinical evidence

Table 1 shows the different types of uncertainties raised by HTA agencies; the total number of uncertainties identified was 342 across the sample, of which trial design (40% of all cases) was the most frequent, followed by the magnitude of the clinical benefit (30%) and the paucity or lack of data (13%). The lack of long-term data on a technology's clinical benefits was a key issue commonly reported by HTA agencies in each assessment. The frequency of uncertainties varied between HTA agencies as follows: 64 cases for PBAC (19%); 62 for each of SMC (18%) and NICE (18%); 42 for HAS (12%); 35 for CADTH and G-BA (10%) ; 22 for TLV (6%) and 20 for IQWIG (6%) (Fig. 2).

Sixty-three percent of the uncertainties raised by NICE (39 cases), 36% by SMC (22) and PBAC (23), 37% by CADTH (12), and 25% by TLV (6) were addressed through various means, such as the generation of further evidence (24 cases); input from clinical, patient or other expertise (23 cases); or they were deemed acceptable (55 cases) in the sense that they did not influence the final recommendation.

Significant variations existed at medicine-indication pair level including how uncertainties were addressed to support the HTA process across countries. In the assessment of alemtuzumab, NICE and SMC included the same phase II trial (randomized and rater-blinded study [CAMMS2233]) to obtain data on effectiveness, safety and HRQoL, used the same source of clinical evidence to perform a meta-analysis and compare alemtuzumab with other DMTs for active RRMS. The inclusion of this meta-analysis led NICE and SMC to highlight several clinical uncertainties around the trial design and clinical benefits. Each emerging issue was addressed by the appropriate NICE committee via consultation with clinical experts and the outcome was a positive recommendation; SMC, however, did not address any of the uncertainties in their reporting. Nevertheless, the different strategies in addressing clinical uncertainties did not seem to have an impact on the final recommendations.

3.5. Economic evidence

A total of 40 reports contained economic models in five countries (submissions to NICE, SMC, TLV, CADTH and PBAC); cost-utility analysis was used most often (70%, n = 28), cost-minimization (23%; n = 9), cost of treatment analysis (3% n = 2) and cost benefit (3%, n = 1). NICE considered cost-utility analyses only, whereas PBAC, TLV and SMC also reported cost-minimization analysis in four (out of 8), three (out of 8) and two (out of 8) cases respectively. CADTH, TLV and PBAC reported one cost analysis each, however for different medicines (respectively, peginterferon, daclizumab and

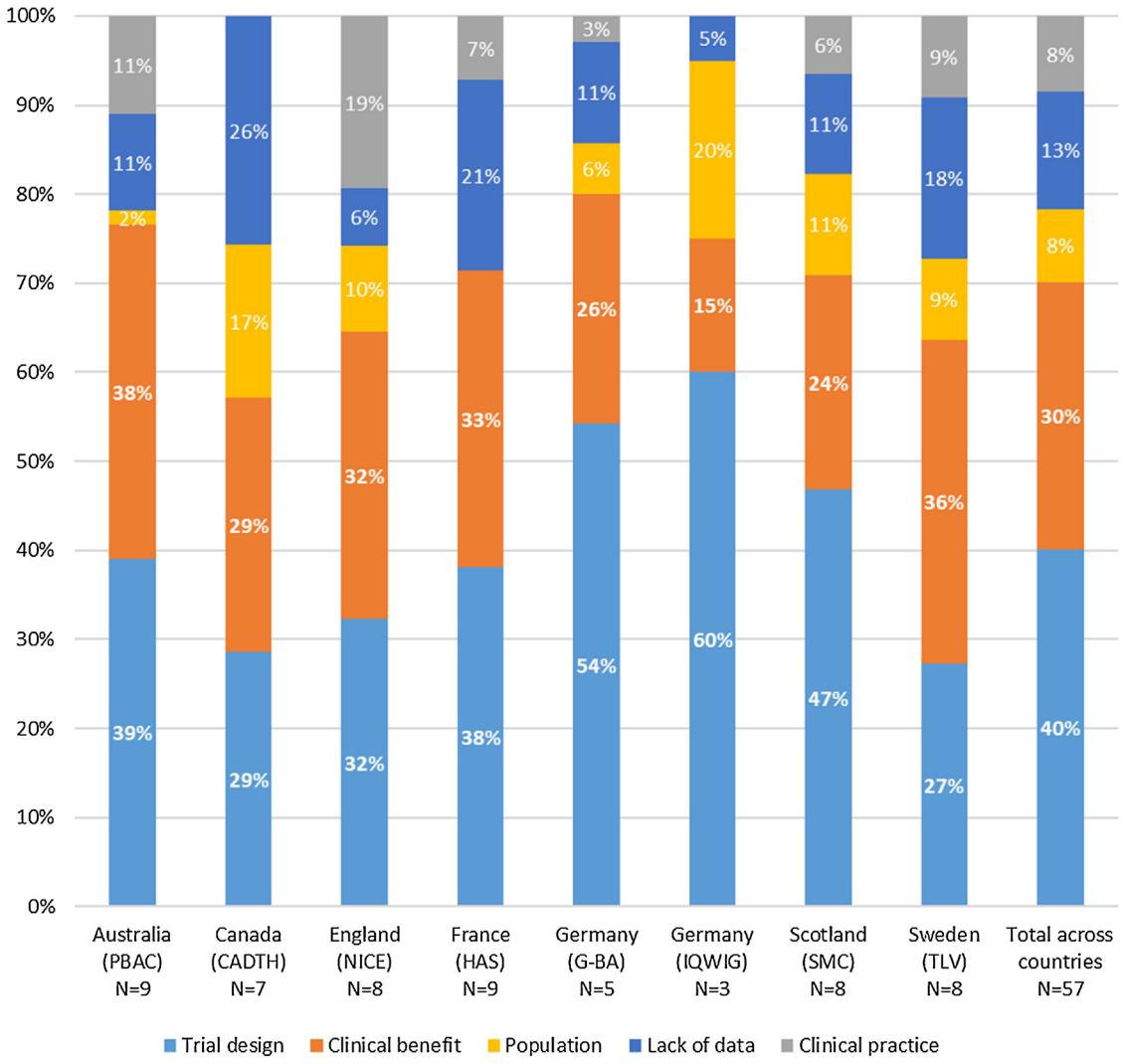

Fig. 2. Type of clinical uncertainties raised and their frequency in HTA reports across seven OECD countries.
CADTH: Canadian Agency for Drugs and Technologies in Health; HAS: Haute Autorité de Santé; IQWIG: Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen; G-BA: Gemeinsamer Bundesausschuss; NICE: National Institute for Health and Care Excellence (NICE); PBAC: Pharmaceutical Benefits Advisory Committee; SMC: Scottish Medicines Consortium; TLV: Dental and Pharmaceutical Benefits Board; N indicates the number of DMT for RRMS assessed by HTA bodies in each of the study countries. Countries are shown in alphabetical order.

Source: Authors' compilations from clinical evidence presented in HTA reports.

teriflunomide). In the G-BA reports the costs for each therapy and comparator therapy were presented without any further explicit analysis.

In each economic analysis more than one medicines were used as a comparator and a total of 74 comparators were considered across the medicine-indication pairs and countries studied. The most frequent type of comparator used was another DMT treatment (55% of cases) and interferons (42% of cases); only in two cases (IFNβ 1b SC assessed by PBAC and natalizumab assessed by CADTH) was 'no therapy' used as comparator. Three agencies also reported budget impact analyses that covered all medicine-indication pairs under consideration (PBAC and SMC) or one singular pair (TLV for teriflunomide).

There were differences in the economic models submitted to HTA agencies in four medicine-indication pairs out of nine, which may have had an impact on the HTA recommendation or its direction. It is not uncommon for different models to lead to different uncertainties; However, even with similar models submitted, different uncertainties where raised, suggesting that differences are not only due to evidence preferences but also to different approaches in assessing the same evidence. For instance, the economic models considered across agencies for fingolimod varied significantly: from cost-minimization analysis compared to natalizumab (TLV and SMC) to cost-utility analysis in comparison with natalizumab (NICE and PBAC), rendering fingolimod cost-effective, to cost-effectiveness analysis comparing fingolimod with interferons (CADTH).

Although TLV and SMC had a similar cost-minimization model in their respective HTA reports, the former reported fingolimod to be cost effective with no restriction, whereas the latter accepted cost-effectiveness only after the submission of a PAS (simple discount) [21]. A different model was presented to PBAC, where, based on economic evidence, it was argued that fingolimod was not cost-effective (the ICER was beyond the suggested thresholds) and deferred their recommendation pending further negotiation with the sponsor, which resulted in a price reduction rendering the ICER acceptable and recommending fingolimod out-of-session based on prior authorisation from PBS. CADTH assessed an economic model with a cost-neutral impact (or a slightly lower price) and deemed this result to be acceptable.

In the case of teriflunomide, manufacturers submitted a cost-minimization analyses to SMC and TLV whereas NICE and CADTH

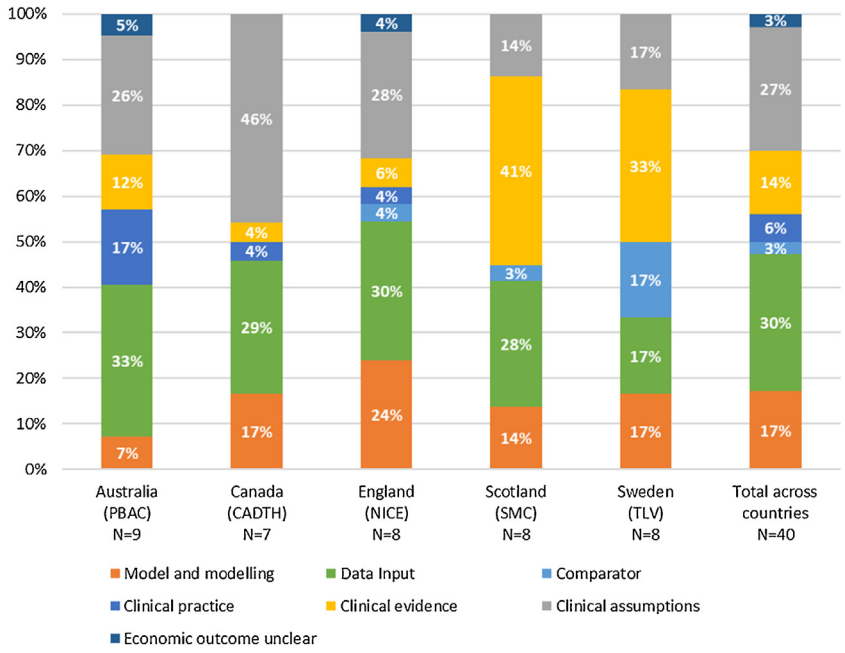

Fig. 3. Type of economic uncertainties raised and their frequency in the HTA reports across seven OECD countries.
CADTH: Canadian Agency for Drugs and Technologies in Health; HAS: Haute Autorité de Santé; IQWIG: Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen; NICE: National Institute for Health and Care Excellence (NICE); PBAC: Pharmaceutical Benefits Advisory Committee; SMC: Scottish Medicines Consortium; TLV: Dental and Pharmaceutical Benefits Board; N indicates the number of DMT for RRMS assessed by HTA bodies in each of the study countries. Countries are shown in alphabetical order.
Source: Authors' compilations of economic evidence presented in HTA reports.

considered similar cost-utility models. TLV and SMC considered different comparators in their analyses (the former used interferon 1-b whereas the latter considered both beta-interferon and glatiramer). Different economic models provided contrasting outcomes to include either listed with restriction for TLV, SMC and NICE or rejected for CADTH.

3.6. Assessment of the economic evidence

The assessment of the economic evidence resulted in several uncertainties (180 cases) raised across the different HTA agencies as follows: 44% (79 cases) by NICE, 24% (42) by PBAC, 16% (29) by SMC, 13% (24) by CADTH and 3% (6) by TLV.

The most frequently raised uncertainties across countries related to data input into each model including utility values, costs and treatment effects used in the model (30%, 54 cases), followed by issues around treatments' clinical assumptions (27%, (49 cases), including assumptions on the long-term effects of medicines, retreatment rates, the discontinuation of treatment, and the return to a rate of progression equivalent to the natural history of the disease Fig. 3. Some variation occurred across agencies: NICE raised a significant number of concerns around the type of model and the associated modeling used (e.g. unsuitability of cost-minimization approach, 19 times); SMC raised concerns about the clinical evidence on which the model was developed (e.g. the indirect comparison used for the clinical data in the economic model, 22 times); and PBAC and CADTH had reservations about the clinical assumptions made in the economic modelling (e.g. the length of the natural course of the disease, 11 times each).

In some cases, the different assessments of the economic evidence had an impact on the final recommendation, although the rationale very often differed. When appraising dimethyl fumarate, SMC and NICE suggested that the medicine would be cost-effective only with a PAS (simple discount) [21]. Although TLV considered a similar cost-utility model and presented equivalent cost-effectiveness outcomes compared with SMC and NICE, the motivation behind their final recommendation differed. TLV stated that, despite the uncertainties around disease progression, the reduction in the frequency of relapses and disease progression was greater than the comparator (interferon b-1-Extavia), thus resulting in lower costs across Expanded Disability Status Scale (EDSS) stages and for relapses. The reduction in costs for the long-term consequences of the disease compensated for the higher cost of the medicine and the additional treatment required. However, TLV reported significant concerns around the benefit of the medicine on disease progression such that a positive recommendation would be conditional on submission of additional documentation to ensure cost-effectiveness of dimethyl fumarate over the long-term, resulting in a LWC recommendation through coverage with evidence development (CED). Similarly, in the case of fingolimod, the requirement of a PAS (simple discount) [21] by NICE and SMC has allowed for a positive recommendation (LWC) on the medicine. In 2011, the manufacturer of fingolimod submitted the first request for evaluation by NICE that rejected fingolimod approving it only once the manufacturer presented a revised analysis with a PAS (simple discount) [21].

3.7. Stakeholder input

Input from external stakeholders played an important role in shaping final recommendations in the majority of the HTA agencies (5 out 7 agencies). Except for of the HTA reports produced by HAS, G-BA and IQWiG where stakeholders such as patients and clinicians are members of the committee and their input is integrated in the process, input from external stakeholders was considered by all other HTA agencies and was stated on all their HTA reports (a total of 41 reports). Looking at the frequencies of stakeholder input, it included input from clinical experts (69% of all cases), patient experts (27%) and other stakeholders, such as caregivers and health economists (4%). Clinician opinion was taken into account by NICE (for all 8 drug-indication pairs), SMC (8), CADTH (7) and PBAC (3). In all of the 9 medicines appraised by SMC, CADTH and PBAC, a

submission from a patient interest group was included, however it was not clear how these submissions influenced the assessment or the final outcome.

3.8. The importance of social value judgments (SVJ)

Upon examination of the overall HTA reporting in our database, a total of 341 SVJs were identified (59% relating to disease and 41% to treatment characteristics) and grouped in 12 categories shown in Table 2. NICE considered the highest number of SVJs (95) followed by HAS (70), SMC (69) and CADTH (61). TLV and PBAC considered 26 and 20 SVJs respectively, whereas G-BA and IQWIG did not consider any SVJs explicitly.

The most common disease-related characteristics raised by all agencies included: the nature of the disease affecting the patient and their surroundings (19%, 66 times) and the impact on patients' everyday activities (19%, 66) such as the ability to work and the autonomy in carrying out normal activities. Commonly raised treatment-related SVJs were highlighted for unmet need for treatments (14%, 48) and the advantages in the administration mode and the frequency of taking the medicines under review (11%; 37). In the latter case the explicit consideration of SVJs had an impact on the final recommendation (in 8 out of 56 reports). For instance, the innovation offered by oral administration in comparison to the standard care in MS, usually administered intravenously, was considered as easing factor of the low clinical benefit identified in the trials, leading to a recommendation of minor improvement in actual benefit (ASMR level IV in France) for fingolimod. NICE, SMC and CADTH considered the oral administration of the medicine as an positive factor in their recommendations for dymethyl fumarate, fingolimod, and teriflunomide. Only NICE, however, explicitly devoted a section of the deliberations for each medicine in considering the presence of SVJs and developed guidance to identify and include them in appraisals [22].

4. Discussion

By using a conceptual framework enabling the study of submitted evidence and its interpretation by competent authorities in 7 OECD countries, we have analysed HTA outputs and final recommendations) on RRMS DMTs and provided an analysis of the similarities and differences in HTA decision-making that may have led to differences in coverage recommendations across seven countries. In evaluating MS treatments, where the burden of disability, the chronic nature and the onset at young age are key features, a transparent and comprehensive HTA process appears to be key in delivering more efficient use of healthcare resources, together with a more equitable access to new and effective technologies.

Key findings across a sample of nine DMTs showed differences in recommendations across HTA agencies, with completely opposite recommendations in three medicine-indication pairs and discordant (but in the same direction and positive) recommendations in five medicine-indication pairs. The evidence collected suggests that the overall heterogeneity in decision-making across HTAs may be correlated with concomitant reasons stemming from differences in the type of clinical and economic evidence considered, with the specific role played by HTA processes in the study countries and their respective preferences for evidence. For example, IQWIG and G-BA openly prefer the inclusion of head-to-head studies to support the clinical benefit of new medicines, whereas clinical benefit data stemming from placebo-controlled trials or indirect comparisons are accepted only if they fulfill rigorous scientific standards; for example, non-adjusted indirect comparisons are considered unsuitable [23]. This is confirmed by our results, where 54% of all uncertainties raised by G-BA were related to the design of the clinical evidence and the associated indirect comparisons presented by the manufacturer. By contrast, placebo-controlled trials can be accepted by other HTA agencies but may carry less gravitas in appraisals; based on the above, the outcome of those assessments is often coverage recommendations running in the opposite direction, as in the case of dimethyl fumarate. This confirms evidence from other studies in different disease areas [14].

The study results have shown heterogeneity in the inclusion or acceptance of clinical evidence across HTA agencies, specifically in the inclusion of indirect comparisons and observational studies, leading in some cases to differences in the assessment of clinical benefit. Although indirect analysis did help mitigate the lack of evidence, the robustness of such an approach may be questionable if not accurately designed [24]. HTA bodies frequently acknowledge the inherent uncertainty of indirect comparisons and raise methodological concerns around their modelling. In many cases in the sample, basic principles such as the similarity of treatment effects or the homogeneity of populations across trials included are still violated in the models designed by manufaturers. In daclizumab, the manufacturer presented various Mixed Treatment Comparisons (MTC) comparing daclizumab versus a range of DMTs, and all HTA agencies raised concerns around the modelling of indirect comparisons due to the significant heterogeneity in the trials used. The statistical methods adopted in indirect comparisons can be a further source of contestability in submitted economic models leading to significant uncertainties regarding the data input for the cost-effectiveness model. Therefore, it appears necessary to develop clear guidelines for undertaking and reporting indirect comparisons in order to ensure their quality and respecting the three basic principles of indirect comparisons (similarity, homogeneity, and consistency). Evidence confirms that diverging HTA recommendations can be a result of variability in the evidence included in the appraisal, heterogeneity in the interpretation of included evidence and in the options that different HTA agencies adopt in dealing with uncertainties [25].

Significant differences were also identified in the detail and the level of analysis when reporting the same clinical evidence. The inclusion of patient subgroup analyses stratified by age or by number of previous therapies (limited to NICE and HAS) could be explained as an attempt to enable access to expensive innovative medicines to those subgroups in greater need for whom the benefits would be greater and the ICER would be acceptable. Often, however, subgroup assessments may be challenging and the resulting treatment effect may be over-estimated due to smaller sample sizes [13]. In these circumstances, additional or real world evidence would constitute a suitable remedy, but this latter option was used rather sparingly in the MS context (three observational studies submitted to CADTH, two to NICE and one each to TLV and SMC) with the exception of France, which considered 17 observational studies across the entire sample.

The high level of heterogeneity between subgroup analyses and the number of trials assessed across countries translated into disparities in the number of endpoints studied. Our results showed a low number of HRQOL factors reported with only 10% of HRQOL considered in the entire sample (n = 68 endpoints), highlighting the need for greater attention on health gain and quality of life data across HTA agencies, particularly for chronic diseases such as MS with a profound impact on patients' and their families' [26–29].

Economic evidence plays a central role in granting access to medicines 5 of the 7 study countries. Equally, the possibility of offering or negotiating a risk sharing agreement as part of the overall HTA process seems to increase the probability of a positive coverage recommendation, albeit with restrictions. The inclusion of a financial PAS lowering the ICER in England and Scotland could reverse a recommendation of NICE and SMC from DNL to LWC [18]). A certain degree of heterogeneity was found in economic models

confirming that, in certain instances, different types of economic model, had a different impact on the HTA outcome, leading to different recommendations as was shown in the case of fingolimod. Moving forward, it might be desirable for HTA agencies to arrive at a consensus on the criteria to be applied in economic models, including: standard level of acceptability for economic studies, common economic comparators, and the inclusion of mitigating factors or criteria to apply consistently in the evaluation of medicines for diseases with a high societal impact such as MS. Specifically, due to the chronic nature of MS, it is important that economic models, accurately capture the long-term costs and effects associated with the disease, including indirect costs. Despite clinical trials producing the best available evidence, and due to stringent selection criteria, there is a danger that results derived from trial populations may not be reflective of the real world MS population. Consequently, those who produce economic models should acknowledge these discrepancies in their projections of average economic costs. The EUnetHTA Core Model®, despite criticisms for its lack of transferability [30], has proposed a set of attributes and methods which could be used in HTA decision frameworks and it may be beneficial to align these attributes and methods across existing HTA frameworks. For cost-effectiveness-driven HTA frameworks, these methods could help modulate ICER thresholds, leaving room to national agencies to automatically arrive at prices for the assessed interventions. In comparative clinical benefit assessment-driven HTA frameworks, standard methods would allow to have a consistent and accepted understanding of additional clinical benefit, allowing the product to be reimbursed subject to negotiation on price.

The high costs and, often, marginal clinical benefits of DMTs in the short-term may lead to a non-comprehensive evaluation of the real magnitude of their benefits, specifically in the long-term [23]. This also raises the question about the possibility of granting special status to medicines for chronic diseases due to the possible impact that these could have on quality and length of life of patients as well as more broadly. Various value frameworks have been developed recently, with emphasis in oncology [31], to address this particular disease area, however other diseases with a large overall budget impact are still in need of more refined HTA tools that capture short as well as long-term effects.

A wide array of considerations and assessments were found relating to the clinical and economic evidence considered in HTAs that are not included in a homogenous way across countries. A different degree of detail in raising and addressing both, clinical and economic evidence concerns was seen with NICE, SMC and PBAC raising a higher number of concerns around the evidence than other agencies. Homogeneity was found in the main type of uncertainties raised, trial design for clinical uncertainties and data input for economic uncertainties, highlighting that agencies look at the same type of key evidence shortfalls. However, significant differences were found elsewhere with concerns being highlighted in some cases, but were completely overlooked in others, for example concerns relating to the model and modelling of economic uncertainties. Our results suggest that it might be beneficial to develop some consensus and, probably, guidance for the evaluation and acceptability of clinical and economic evidence necessary to assess the clinical and cost effectiveness consistently across settings, confirming findings elsewhere [15,25].

Our results show common concerns around the absence of long-term effectiveness data for DMTs and the relative absence of real world evidence (RWE) generation, with the exception of dimethyl fumarate and glatiramer acetate, and exclusive to Sweden, where the decision was restricted in the first assessment pending the generation of effectiveness evidence through CED. Improving the observational data collection in composite registries can inform ongoing HTA processes and potentially facilitate earlier use of DMTs and needs to be strengthened; the latter has been found to reduce accumulation of irreversible long-term damage and decrease the high socioeconomic burden of the disease [32].

Evidence from the literature confirmed that stakeholders should be actively engaged by those conducting HTAs to understand and account for their perspectives at various stages of the HTA process and to gather evidence on key value dimensions of medicines that otherwise might be overlooked [33,34]. Yet, our study showed that evidence from clinicians, patients and caregivers is routinely and explicitly considered in some settings (NICE, CADTH and SMC), whereas such input is considered in a much less systematic and more ad hoc way (PBAC and TLV) or is not considered at all (HAS and IQWIG). Assuming that stakeholder input is important in making recommendations that are all-inclusive, it would be desirable to identify better ways to include standardized stakeholder input in HTA decision-making.

SVJs are used in varying degrees across countries to inform decision-making. This is compatible with evidence at system level [35], however, heterogeneity was found in their consideration, which, in some cases, clearly impacted the final recommendation. In order to capture the different dimensions of value of medicines that have a significant impact on patients and society it is, therefore, important and equitable to include different elements of value that go beyond clinical and cost-effectiveness. In the context of MS, several studies have indicated that early treatment with DMTs can delay the development of new symptoms [23–28] whereas treatment delays could lead to the development of severe and irreversible neurological disability [29,30] leading to higher disease-management costs. Oncology and rare disease-specific HTAs seem to be capturing the impact of disease-specific dimensions (rarity, unmet clinical need, end-of-life criteria, among others) in better determining willingness to pay [36,37]. The inclusion of these elements allows the prioritization of some treatments over others, despite their high cost. Although some SVJs seem to have been included in the assessment of RRMS DMTs, the impact on families and caregivers is not routinely considered despite its significance.

Our analysis is not without limitations. First, with the exception of the German G-BA and the Swedish TLV, which are regulatory bodies, all other HTA agencies have an advisory role, which practically means that their recommendations do not constitute coverage decisions. However, in England, Scotland, Australia and France, these recommendations are legally binding and, consequently, are implemented by the respective competent authorities on coverage. In Canada, CADTH's recommendations are taken into account by the provincial drug benefit plans, who are the final decision makers. The letter may or may not endorse these recommendations when it comes to making a coverage decision, although, evidence suggests that CADTH's HTA recommendations are usually translated into coverage decisions at provincial level [38]. Second, there may be variable levels of detail in HTA reports across countries and all the information and discussions that were associated with the final recommendation may not be included. However, the analysis, the main reasons and the rationale for final recommendations were clearly identified in all reports offering a good understanding of the different decision-making processes. Finally, whereas information about the context within which recommendations were arrived at may not have been captured, the operating assumption in our research has been that HTA processes have been transparent and that publicly available reports adequately capture process, rationale and decision criteria.

5. Conclusion

In this study, we identified significant heterogeneity in evidence preferences across countries, including the acceptance of clinical and economic evidence across settings and, specifically,

the evidence from indirect comparisons, the use of real world evidence and the type of economic model considered. Health care systems operate in an ever challenging and constrained resource environment and observe demand for and cost of new technologies increase steadily. As the pursuit of efficiency and budget impact have become key in the assessment of technologies, improvements in HTA processes are necessary that would enable more equitable access to technologies across countries, together with a more efficient use of healthcare resources. To do so, appropriate guidelines and systematic approaches as well as some consensus need to be developed; this is particularly relevant for evidence acceptability, the inclusion of input from different stakeholders, the incorporation of social value judgements in a more systematic, rather than ad hoc way, and the wider incorporation of RWE in what appears to be a continuous re-assemment of technologies, particularly in disease areas with a high social impact such as MS. Our study has highlighted significant variations in all the above areas that may be desirable to address in the near future.

Declarations of interest

The development of the article was supported partly by an unrestricted educational grant from Roche.

We confirm that the manuscript has been read and approved by all named authors and that there are no other persons who satisfied the criteria for authorship but are not listed. We further confirm that the order of authors listed in the manuscript has been approved by all of us.

We confirm that we have given due consideration to the protection of intellectual property associated with this work and that there are no impediments to publication, including the timing of publication, with respect to intellectual property. In so doing we confirm that we have followed the regulations of our institutions concerning intellectual property.

We understand that the Corresponding Author is the sole contact for the Editorial process. She is responsible for communicating with the other authors about progress, submissions of revisions and final approval of proofs. We confirm that we have provided a current, correct email address which is accessible by the Corresponding Author and which has been configured to accept email from e.visintin@lse.ac.uk.

Acknowledgements

We are grateful to the following experts for providing useful feedback: Olina Efthymiadou, Jean Mossman, Anna-Maria Fontrier and Federico Grimaccia.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.healthpol.2018.08.019.

References

[1] Dutta R, Trapp BD. Pathogenesis of axonal and neuronal damage in multiple sclerosis. Neurology 2007;68:S43–54.
[2] Baneke P, Browne P, Thompson AJ, Taylor B, Battaglia M, Pandit L, et al. MSIF atlas of MS database update: Multiple sclerosis resources in the world 2013. Multiple Sclerosis Journal 2013;19:652.
[3] Ernstsson O, Gyllensten H, Alexanderson K, Tinghog P, Friberg E, Norlund A. Cost of illness of multiple sclerosis—a systematic review. PLoS One 2016;11:e0159129.
[4] Weber MS, Menge T, Lehmann-Horn K, Kronsbein HC, Zettl U, Sellner J, et al. Current treatment strategies for multiple sclerosis-efficacy versus neurological adverse effects. Current Pharmaceutical Design 2012;18:209–19.
[5] Paty DW, Li DK. Interferon beta-1b is effective in relapsing-remitting multiple sclerosis. II. MRI analysis results of a multicenter, randomized, double-blind, placebo-controlled trial. UBC MS/MRI Study Group and the IFNB Multiple Sclerosis Study Group. Neurology 1993;43:662–7.
[6] Johnson KP, Brooks BR, Cohen JA, Ford CC, Goldstein J, Lisak RP, et al. Copolymer 1 reduces relapse rate and improves disability in relapsing-remitting multiple sclerosis: results of a phase III multicenter, double-blind placebo-controlled trial. The Copolymer 1 Multiple Sclerosis Study Group. Neurology 1995;45:1268–76.
[7] Jacobs LD, Cookfair DL, Rudick RA, Herndon RM, Richert JR, Salazar AM, et al. Intramuscular interferon beta-1a for disease progression in relapsing multiple sclerosis. The Multiple Sclerosis Collaborative Research Group (MSCRG). Annals of Neurology 1996;39:285–94.
[8] Kappos L, Antel J, Comi G, Montalban X, O'Connor P, Polman CH, et al. Oral fingolimod (FTY720) for relapsing multiple sclerosis. The New England Journal of Medicine 2006;355:1124–40.
[9] O'Connor P, Wolinsky JS, Confavreux C, Comi G, Kappos L, Olsson TP, et al. Randomized trial of oral teriflunomide for relapsing multiple sclerosis. The New England Journal of Medicine 2011;365:1293–303.
[10] Kappos L, Gold R, Miller DH, Macmanus DG, Havrdova E, Limmroth V, et al. Efficacy and safety of oral fumarate in patients with relapsing-remitting multiple sclerosis: a multicentre, randomised, double-blind, placebo-controlled phase IIb study. Lancet 2008;372:1463–72.
[11] Johnson KP, Brooks BR, Ford CC, Goodman AD, Lisak RP, Myers LW, et al. Glatiramer acetate (Copaxone): comparison of continuous versus delayed therapy in a six-year organized multiple sclerosis trial. Multiple Sclerosis 2003;9:585–91.
[12] Kappos L, Freedman MS, Polman CH, Edan G, Hartung HP, Miller DH, et al. Effect of early versus delayed interferon beta-1b treatment on disability after a first clinical event suggestive of multiple sclerosis: a 3-year follow-up analysis of the BENEFIT study. Lancet 2007;370:389–97.
[13] Comi G. Clinically isolated syndrome: the rationale for early treatment. Nature Clinical Practice Neurology 2008;4:234–5.
[14] Kanavos PT, Tinelli M, Efthymiadou O, Visintin E, Grimaccia F, Mossman J. Towards better outcomes in multiple sclerosis by addressing policy change. London: London school of economics and Political science; 2016.
[15] Nicod E, Kanavos P. Commonalities and differences in HTA outcomes: a comparative analysis of five countries and implications for coverage decisions. Health Policy 2012;108:167–77.
[16] Fischer KE, Heisser T, Stargardt T. Health benefit assessment of pharmaceuticals: an international comparison of decisions from Germany, England, Scotland and Australia. Health Policy 2016;120:1115–22.
[17] Nicod E. Why do health technology assessment coverage recommendations for the same drugs differ across settings? Applying a mixed methods framework to systematically compare orphan drug decisions in four European countries. The European Journal of Health Economics 2017;18:715–30.
[18] Spinner DS, Birt J, Walter JW, Bowman L, Mauskopf J, Drummond MF, et al. Do different clinical evidence bases lead to discordant health-technology assessment decisions? An in-depth case series across three jurisdictions. ClinicoEconomics and Outcomes Research 2013;5:69–85.
[19] O'Donnell JC, Pham SV, Pashos CL, Miller DW, Smith MD. Health technology assessment: lessons learned from around the world—an overview. Value Health 2009;12(Suppl. 2):S1–5.
[20] Schwarzer R, Siebert U. Methods, procedures, and contextual characteristics of health technology assessment and health policy decision making: comparison of health technology assessment agencies in Germany, United Kingdom, France, and Sweden. International Journal of Technology Assessment in Health Care 2009;25:305–14.
[21] National Institute for Health and Care Excellence (NICE). List of recommended technologies that include a commercial arrangement; 2018.
[22] National Institute for Health and Care Excellence (NICE). What are the societal values that need to be considered when making decisions about trade-offs between equity and efficiency? London; 2014.
[23] Ivandic V. Requirements for benefit assessment in Germany and England—overview and comparison. Health Economics Review 2014;4:12.
[24] Hoaglin DC, Hawkins N, Jansen JP, Scott DA, Itzler R, Cappelleri JC, et al. Conducting indirect-treatment-comparison and network-meta-analysis studies: report of the ISPOR Task Force on Indirect Treatment Comparisons Good research practices: part 2. Value Health 2011;14:429–37.
[25] Nicod E. Why are there differences in hta recommendations across countries? A systematic comparison of hta decision processes for a sample of orphan drugs in four countries. Value Health 2014;17:A540.
[26] Salter AR, Cutter GR, Tyry T, Marrie RA, Vollmer T. Impact of loss of mobility on instrumental activities of daily living and socioeconomic status in patients with MS. Current Medical Research and Opinion 2010;26:493–500.
[27] Kobelt G, Berg J, Atherly D, Hadjimichael O. Costs and quality of life in multiple sclerosis: a cross-sectional study in the United States. Neurology 2006;66:1696–702.
[28] Kobelt G, Berg J, Lindgren P, Fredrikson S, Jonsson B. Costs and quality of life of patients with multiple sclerosis in Europe. Journal of Neurology, Neurosurgery, and Psychiatry 2006;77:918–26.
[29] Stawowczyk E, Malinowski KP, Kawalec P, Mocko P. The indirect costs of multiple sclerosis: systematic review and meta-analysis. Expert Review of Pharmacoeconomics & Outcomes Research 2015;15:759–86.
[30] Oortwijn W, Broos P, Vondeling H, Banta D, Todorova L. Mapping of health technology assessment in selected countries. International Journal of Technology Assessment in Health Care 2013;29:424–34.

[31] Schnipper LE, Davidson NE, Wollins DS, Tyne C, Blayney DW, Blum D, et al. American Society of Clinical Oncology Statement: a conceptual framework to assess the value of cancer treatment options. Journal of Clinical Oncology 2015;33:2563–77.
[32] Giovannoni G, Butzkueven H, Dhib-Jalbut S, Hobart J, Kobelt G, Pepper G, et al. Brain health: time matters in multiple sclerosis. Multiple Sclerosis and Related Disorders 2016;9(Suppl. 1):S5–48.
[33] Cavazza M, Jommi C. Stakeholders involvement by HTA organisations: why is so different? Health Policy 2012;105:236–45.
[34] Gagnon MP, Candas B, Desmartis M, Gagnon J, La Roche D, Rhainds M, et al. Involving patient in the early stages of health technology assessment (HTA): a study protocol. BMC Health Services Research 2014;14:273.
[35] Angelis A, Lange A, Kanavos P. Using health technology assessment to assess the value of new medicines: results of a systematic review and expert consultation across eight European countries. European Journal of Health Economics 2017;19(1):123–52.
[36] Schwarzer R, Rochau U, Saverno K, Jahn B, Bornschein B, Muehlberger N, et al. Systematic overview of cost-effectiveness thresholds in ten countries across four continents. Journal of Comparative Effectiveness Research 2015;4:485–504.
[37] Chabot I, Rocchi A. Oncology drug health technology assessment recommendations: canadian versus UK experiences. ClinicoEconomics & Outcomes Research 2014;6:357–67.
[38] Allen Nicola, Walker Stuart R, Liberti Lawrence, Sehgal Chander, Sam Salek M. Evaluating alignment between Canadian Common Drug Review reimbursement recommendations and provincial drug plan listing decisions: an exploratory study. CMAJ Open 2016;4:E674–8, http://dx.doi.org/10.9778/cmajo.20160006, published online November 4.