VALUE IN HEALTH ▮ (2017) ▮▮▮–▮▮▮

Available online at www.sciencedirect.com

ScienceDirect

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

Dealing with Uncertainty and Accounting for Social Value Judgments in Assessments of Orphan Drugs: Evidence from Four European Countries

Elena Nicod, PhD[1,2,*], *Karen Berg Brigham, JD, MPH*[3,4], *Isabelle Durand-Zaleski, MD, PhD*[3,4,5], *Panos Kanavos, PhD*[1]

[1]*Department of Social Policy, LSE Health, London School of Economics and Political Science, London, UK;* [2]*Center for Research on Health and Social Care Management, Bocconi University, Milan, Italy;* [3]*Université Paris Est Créteil Val de Marne (UPEC), Créteil, France;* [4]*URC Eco Ile-de-France (AP-HP), Paris, France;* [5]*ECEVE UMRS 1123, UEC—Hôpital Robert Debré (AP-HP), Paris, France*

ABSTRACT

Objectives: To better understand the reasons for differences in reimbursement decisions for orphan drugs in four European countries that were not readily apparent from health technology assessment (HTA) reports and operating procedures. **Methods:** Semistructured interviews with representatives of HTA bodies in England, Scotland, Sweden, and France were conducted. An interview topic guide was developed on the basis of findings from a systematic comparison of HTA decisions for 10 orphan drugs. Qualitative thematic data analysis was applied to the interview transcripts using the framework approach. **Results:** Eight representatives from the four HTA bodies were interviewed between March and June 2015. Evidentiary requirements and approaches to dealing with imperfect or incomplete evidence were explored, including trial design and duration, study population and subgroups, comparators, and end points. Interviewees agreed that decisions regarding orphan drugs are made in a context of lower quality evidence, and the threshold of acceptable uncertainty varied by country. Some countries imposed higher evidentiary standards for greater clinical claims, which may be more challenging for orphan diseases. The acceptability of surrogate end points was not consistent across countries nor were the validation requirements. The most common social value judgments identified related to innovation, disease severity, and unmet need. Differences were seen in the way these concepts were defined and accounted for across countries. **Conclusions:** Although agreement was seen in evidentiary requirements or preferences, there were subtle differences in the circumstances in which uncertain evidence may be considered acceptable, possibly explaining differences in HTA recommendations across countries.

Keywords: health technology assessment, orphan drugs, rare diseases, value assessments.

Introduction

Health technology assessment (HTA) aims to ensure that technologies offered are safe and efficacious and provide value for money [1]. Although value is often considered within the context of efficiency—reimbursing only the most efficient technologies within an allowable budget—this does not necessarily account for what matters most to patients or to society in general [2]. Indeed, certain aspects of value are difficult to capture and yet may provide benefits to both, such as innovation that results in a direct benefit to patients through improved prognosis or quality of life and also indirect societal benefits in terms of increased productivity and knowledge spillovers.

Despite using the same evidence and similar outcome measures and criteria, HTA assessments of a given drug may lead to contrary results in different countries [3]. This is particularly true with respect to orphan drugs, for which the general rules regarding appropriate evidence may be difficult to apply to small populations facing very serious chronic or life-limiting diseases [4]. Orphan drug trials are often characterized by lower quality evidence compared with nonorphan drugs [5,6]. Moreover, high acquisition costs often result in orphan drugs not being found to be cost-effective [7]. Nonetheless, orphan drugs often undergo the same HTA processes as drugs for more prevalent conditions.

In the face of imperfect evidence and high uncertainty in assessing orphan drugs, HTA bodies may rely on different attributes of value or approaches to dealing with imperfect evidence. Acceptability of uncertainty depends on the tools used to address uncertainty and on the judgment of the decision makers, who may consider additional qualitative criteria, such as disease or treatment characteristics [8].

Understanding the rationales underlying conflicting decisions is challenging. Although the internal regulations of HTA bodies explain the operating framework and the opinions or recommendations document the evidence considered and the basis for the decision, certain subtleties may not be captured even in the most complete documentation. A better understanding is therefore needed about how HTA bodies value orphan drugs and deal with issues of rarity.

* *Address correspondence to*: Elena Nicod, Center for Research on Health and Social Care Management, Bocconi University, Via Roentgen 1, Milan 20126, Italy.

E-mail: elena@nicod.com.

1098-3015$36.00 – see front matter

http://dx.doi.org/10.1016/j.jval.2017.03.005

We previously analyzed the decisions of 4 HTA bodies for 10 orphan drug-indication pairs on the basis of the opinions and in light of each entity's internal regulations [9], and we explored scientific and social value judgments used in the assessment of orphan drugs [10]. A number of reasons for differences in HTA recommendations were identified throughout the decision process and across countries. Building on these findings, this study aimed to develop a broader perspective about how value is assessed for orphan drugs and how differences affect reimbursement decisions on the basis of interviews of representatives of four European HTA bodies.

Methods

Purposeful sampling was used to select the study countries, each of which undertakes assessments using well-established processes and criteria, has publicly available reports, and represents a cross selection in terms of HTA approach and perspective (Table 1). These included the National Institute for Health and Care Excellence (NICE, England), the Scottish Medicines Consortium (SMC, Scotland), the Dental and Pharmaceutical Benefits Board (Tandvårds- och läkemedelsförmånsverket, Sweden), and the French National Authority for Health (*Haute Autorité de Santé* [HAS], France). HTA body representatives from each study country were identified by partners of a European research consortium, Advance-HTA [11]. These HTA bodies have either regulatory or advisory roles, in which their decisions will be automatically implemented in the former and accounted for by the final decision maker in the latter (Table 1). Furthermore, orphan drugs do not have a special status in the study countries, with the exception of SMC, in which greater uncertainty or higher incremental cost-effectiveness ratios (ICERs) may be accepted if the requirements for their modifiers are fulfilled [12].

We conducted semistructured interviews using an interview topic guide developed by the lead author and reviewed by all co-authors (see eAppendix A in Supplemental Materials found at http://dx.doi.org/10.1016/j.jval.2017.03.005). It included open-ended questions derived from actual scenarios that arose in the context of our cross-national comparison of 10 orphan drugs. Interview questions were divided into themes, including 1) the general evidentiary requirements for orphan drugs regarding primary and nonprimary evidence, trial duration, and clinical and surrogate end points; 2) other evidence and considerations around quality-of-life data and qualitative critera (innovation, unmet need, and disease severity) and the consistency in the considerations across decisions; 3) dealing with uncertainty relating to orphan drug characteristics; and 4) stakeholder involvement. An email invitation to participate in a face-to-face or telephonic interview along with the topic guide was sent to each interviewee. Anonymity was assured, and interviews were recorded and transcribed and sent to the interviewees for comment and validation. The study protocol was reviewed pursuant to the London School of Economics Research Ethics procedure and was found to be exempt (see eAppendix B in Supplemental Materials found at http://dx.doi.org/10.1016/j.jval.2017.03.005).

Qualitative thematic data analysis was undertaken using the framework approach [13]. Subthemes within each general theme were identified and inductively coded, and a matrix was created to facilitate comparison of each subtheme across the four HTA bodies. The key findings from each of these subthemes were summarized in tables that incorporated illustrative quotes. The initial findings were discussed among the co-authors, and a list of follow-up questions was developed to complement the interviews in which information was unclear or incomplete. These additional questions were sent to each interviewee along with the summary findings for their particular HTA body for confirmation. Results focused on the contrasts across countries identified within each theme. Themes were reorganized as follows: 1) clinical evidence and uncertainty, 2) comparators, 3) treatment outcomes and safety, and 4) additional qualitative criteria.

Each theme portrays the agencies' perspectives about the clinical evidence appraised and whether evidence for orphan drugs is characterized by greater uncertainty compared with drugs for more prevalent conditions. The evidence base used for HTA is imperfect or incomplete, and therefore uncertain, because it relies on estimated values from experimental or observational studies [14–16]. Decision makers make scientific value judgments about the extent to which uncertain evidence is acceptable. These judgments include whether the evidence presented fully and accurately captures the effect of the intervention, whether it is generalizable to the local context of the decision, whether quality-of-life changes are accurately captured, or whether it is appropriate to impose restrictions to population subgroups [14]. We aimed to obtain additional insights on the appraisal processes in terms of the HTA bodies' approaches to dealing with uncertain evidence, including the circumstances under which imperfect or incomplete evidence that does not accurately capture the effect of the intervention may be deemed acceptable.

Results

Eight representatives from the four HTA bodies were interviewed between March and June 2015. Interviewees occupied senior positions in their agencies (e.g., Head of the Technology Appraisal Programme, Head Economist or Pharmacist, and Chair of the Appraisal Committee). Interviews were conducted face-to-face and, in one case, by telephone, lasting 1 to 3.5 hours. Responses are summarized in Figures 1 and 2 and presented in Table 2,

Table 1 – Study countries, HTA bodies, and types of HTA.

Study country	HTA body	Type of HTA
England	NICE: National Institute for Health and Care Excellence (regulatory body)	Clinical and cost-effectiveness, national health and personal social services perspective
Scotland	SMC: Scottish Medicines Consortium (advisory body to the NHS boards)	Clinical and cost-effectiveness, national health and personal social services perspective
Sweden	TLV: Dental and Pharmaceutical Benefits Board (regulatory body)	Clinical and cost-effectiveness, societal perspective
France	HAS: Haute Autorité de Santé (Comité de la Transparence) (advisory body to the Ministry of Health)	Benefit-risk ratio, clinical benefit driving the coverage rate (SMR), and relative improvement in clinical benefit driving the pricing scheme (ASMR)

ASMR, relative improvement in clinical benefit ("Amélioration du Service Médical Rendu"); HTA, health technology assessment; NHS, National Health Service; SMR, clinical benefit ("Service Médical Rendu").

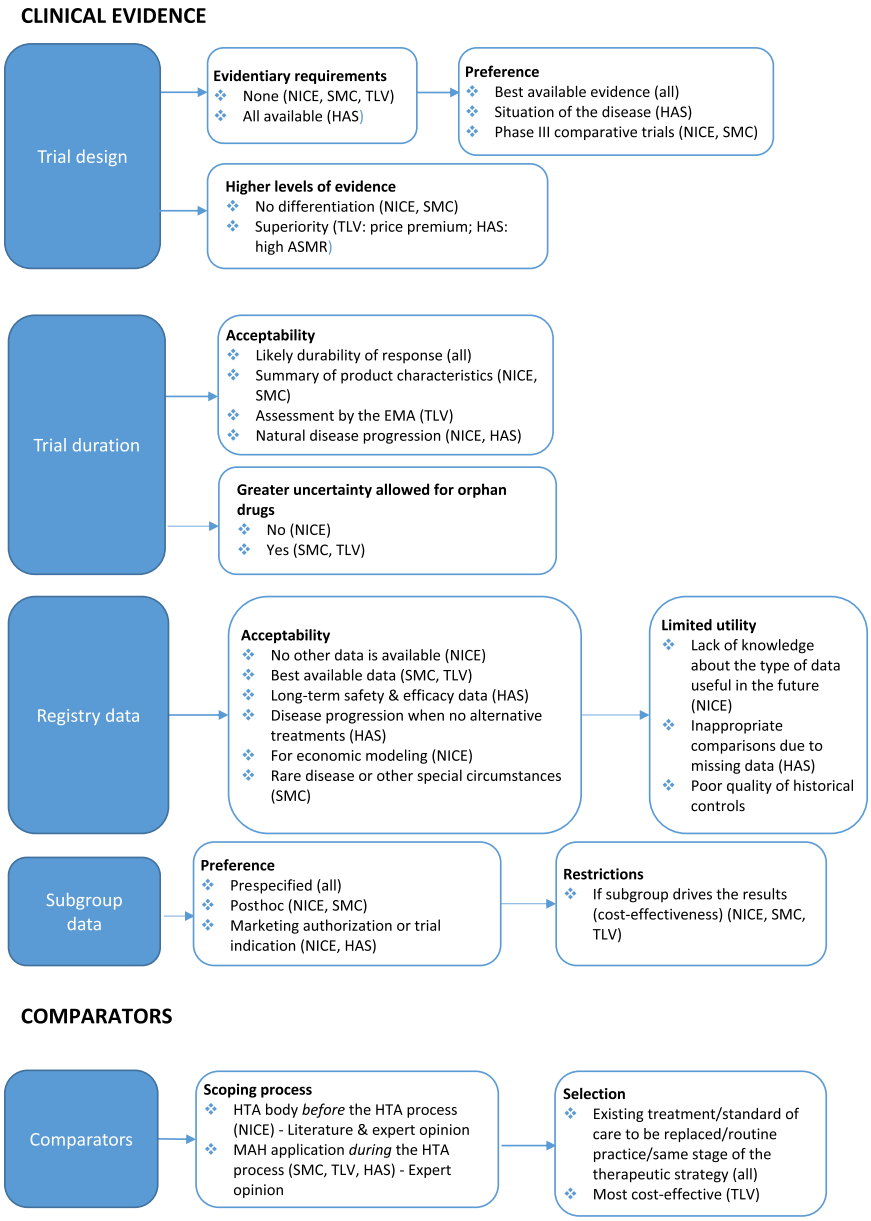

Fig. 1 – Visual representation of interview responses (clinical evidence). ASMR, relative improvement in clinical benefit ("Amélioration du Service Médical Rendu"); EMA, European Medicine Agency; HAS, Haute Autorité de Santé; HTA, health technology assessment; MAH, marketing authorization holder; NICE, National Institute for Health and Care Excellence; SMC, Scottish Medicines Consortium; TLV, Dental and Pharmaceutical Benefits Board.

which includes illustrative quotations.

Clinical Evidence

Trial design

None of the HTA bodies imposes formal requirements regarding minimum levels of evidence for orphan drugs, although phase III comparative trials are preferred. HAS also requires all existing and available data at the time of the HTA submission.

An important distinction was seen in expectations about the quality of the evidence submitted when examined within the context of the clinical claim. The TLV has higher scientific and methodological demands for superior efficacy with a price premium, and greater uncertainty is accepted for noninferior efficacy (and low price) or for treating otherwise untreatable diseases. HAS judges the quality of the evidence according to the situation of the disease in terms of prevalence and recruitable patients, and the highest relative improvement in clinical benefit ("Amélioration du Service Médical Rendu") (ASMR) rating should demonstrate a positive effect on survival.

Trial duration

In considering the appropriate trial duration, all the HTA bodies are concerned about the likely durability of response. NICE and HAS also account for the natural progression of the disease. The TLV looks to the European Medicine Agency's assessment for guidance in this regard, whereas NICE and SMC look to the summary of product characteristics supplied by the marketing authorization holder (MAH). Overall, the SMC, TLV, and HAS were

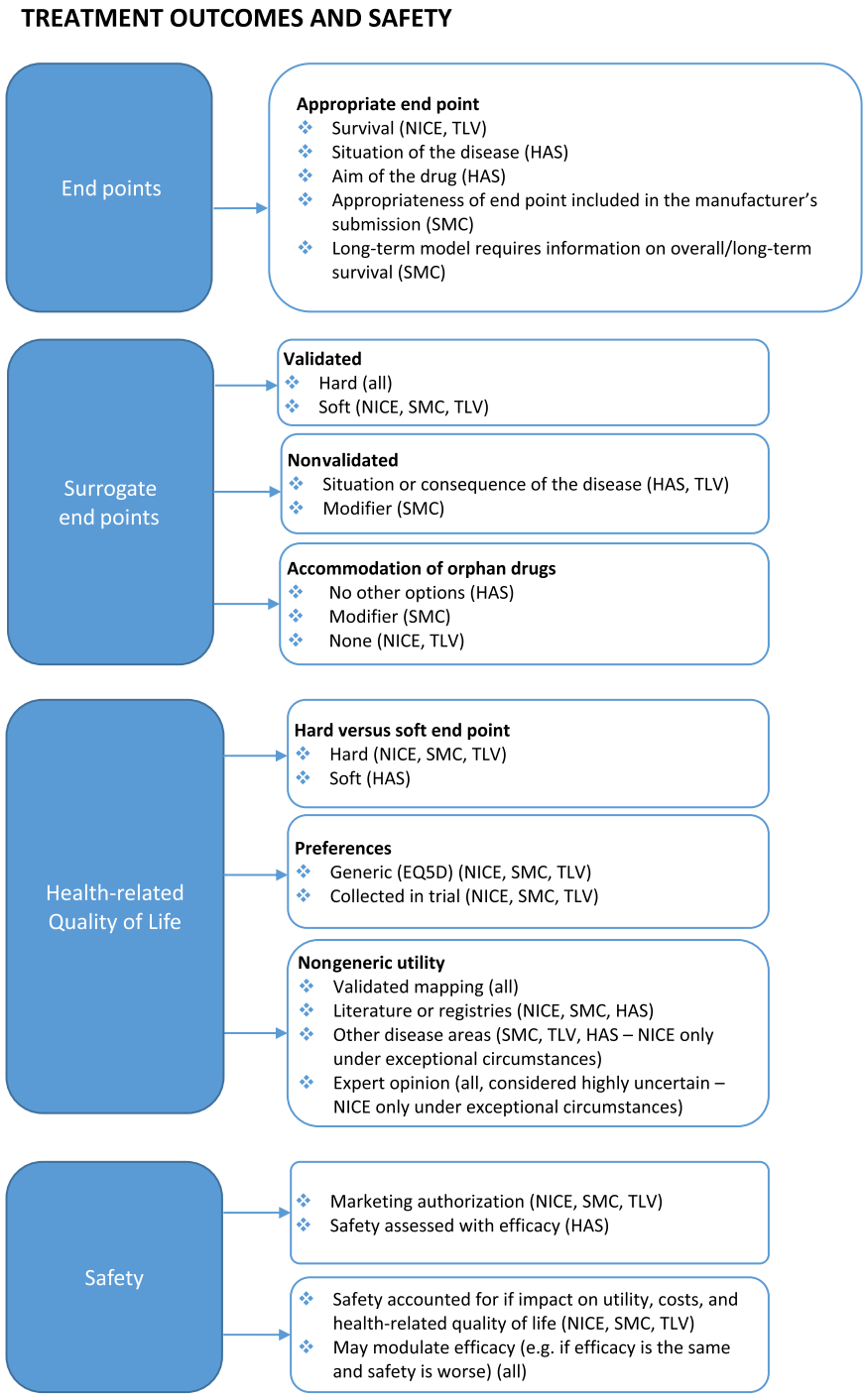

Fig. 2 – Visual representation of interview responses (treatment outcomes and safety). EQ-5D, EuroQol five-dimensional questionnaire; HAS, Haute Autorité de Santé; HTA, health technology assessment; NICE, National Institute for Health and Care Excellence; SMC, Scottish Medicines Consortium; TLV, Dental and Pharmaceutical Benefits Board.

more willing than NICE to accept greater uncertainty on treatment duration for orphan drugs.

Registry data

Although rarely used, the agencies agreed that registry data and historical controls are acceptable when no other data are available (NICE), when these are the best data available (SMC, TLV), to collect long-term data on safety and efficacy or on disease progression when no alternative treatments exist (HAS), for economic modeling purposes (NICE), or when the disease is rare or other special circumstances are seen (SMC). The limited use of historical controls may be explained by the lack of knowledge about the type of data useful in the future (NICE), missing

Table 2 – Summary of interview responses and illustrative quotations.

			NICE	SMC	TLV	HAS	Illustrative quotations/comments
Trial design	Requirements	None	☑	☑	☑		No formal requirements
		All available evidence				☑	HAS: "HAS requires all the clinical trial data available at the time of the assessment"
	Preferences	Best or highest available evidence	☑			☑	HAS: "HAS has a preference for demonstrative data, which means data that are the highest level of evidence (e.g. phase III comparative well-designed and conducted trial)"
		Phase III comparative trials	☑	☑			NICE: "The Committee feels more comfortable about making decisions on clinical effectiveness based on phase III trials, but it is very rare to actually have phase III trials with the correct comparator"
		Requirements similar for all drugs	☑	☑			SMC: "each case is viewed upon its own merits"
		Higher methodological requirements for superior/higher efficacy claim			☑	☑	TLV: "If their price is really low (and clinical claim is non-inferiority), then any uncertainty is ok as long as patients don't die (which has already been checked by the EMA)" HAS: A higher claim should demonstrate in a good way the effect of the treatment—"ASMR I is granted for drugs that have a demonstrated in a good way a substantial effect on survival… The ASMR IV is for a demonstration that is not so perfect and with a quantity of effect which exists but is not very important"
		Lower methodological requirements when the consequence of the decision is severe			☑		TLV: "Greater uncertainty accepted if the consequence of the decision is severe"
		Quality of the evidence is assessed according to the situation of the disease (prevalence and number of recruitable patients)				☑	HAS: "accounts for the real situation of the disease, considering the prevalence and number of patients that are recruitable in trials, as often seen for orphan drugs"
	Historical controls, acceptability criteria	When no other data are available	☑				NICE: "Historical controls are rarely seen mainly in cases when no other data is available"
		When it is the best evidence available		☑	☑		
		When no other treatments are available and to obtain data on disease progression				☑	HAS: "HAS is very much in favour for prospective appropriate data collection on natural history of a disease that can serve as a comparison when another comparison is not possible"
		When the disease is rare or other special circumstances are seen		☑			SMC: "the acceptability of registry data by the Committee would depend on many factors already discussed (e.g. rarity, etc.)…"
	Registry data use	Historical controls	☑	☑	☑	☑	
		Natural progression of the disease (e.g., to obtain long-term data)	☑			☑	NICE: "lifelong modelling of the disease and therefore need long-term data about disease progression, which will never come from any trial"
		Treatment efficacy and safety				☑	HAS: "at the first assessment for reimbursement, in general only short-term data is available and orphan disease are in majority chronic diseases so they also rely on registries to have longer term data on efficacy first, and safety second"
Trial length	Acceptability criteria	Natural progression of the disease	☑			☑	NICE: "the Committee always welcomes data on natural history of the disease to validate any extrapolation curves"

continued on next page

Table 2 – *continued*

			NICE	SMC	TLV	HAS	Illustrative quotations/comments
		Likely durability of treatment response	☑	☑	☑	☑	HAS: "If the duration is too short compared to the natural course of the disease then it will be criticised" NICE: "The Committee always welcomes data on the parameter, or seeks sensitivity analyses with different assumptions if no data is available" SMC: "… the likely durability of treatment response (may be informed by other sources)" HAS: "… sufficiently long to generate solid evidence about the type of benefit to the patient" TLV: "The trial length needs to cover the time span up to the point where we can see that both treatments converge"
		Corresponds to the EMA assessment			☑		TLV: "This is done by the EMA and TLV trusts that the right recommendations were given"
		SPC advice for treatment duration	☑	☑			NICE: "NICE would be bound by the treatment duration specified in the SPC, unless a stopping rule is proposed by the company and supported by the clinical community." SMC: "It will relate to consideration of factors such as the duration of the trial relative to what the SPC advises in terms of treatment duration"
		Greater flexibility accepted for orphan drugs	☒	☑	☑		NICE: "we do not differentiate orphan drugs" SMC: "Trial duration isn't something that is specifically teased out as an issue but may be something that is noted as a general weakness of the evidence base (particularly if very short in relation to a very long-term economic model). To the extent that we offer greater flexibility in dealing with the general limitations with orphan drugs, issues with limitations in trial duration would be afforded similar flexibility." TLV: "greater uncertainty regarding the clinical effect is accepted"
Study population	Subgroup data, acceptability criteria	Prespecified	☑	☑	☑	☑	HAS: "The Transparency Committee will not be confident in the results if the subgroup was post hoc, and have a clear preference for pre-planned or pre-specified subgroups … has to be pre-specified in the protocol of the trial"
		Post hoc	☑	☑	☒		NICE: "sometimes the population in the licence is from a post hoc group, in which case NICE needs to consider it. Otherwise, these are very rarely included, only if there is a strong biological plausibility of a strong cost-effectiveness argument for including it" TLV: "Subgroups must have been pre specified before using them, it is an absolute rule" SMC: "consideration would be given to whether the subgroup was pre-specified or post hoc and also the relative size of the subgroup and the potential significance of any results"
		Relative size of the subgroup	☒	☑			NICE: "ideally, small subgroups are not considered, but that depends on what population the licence covers"

continued on next page

		If it is the only available evidence for a very severe rare condition, nonspecified data could be acceptable			☑		
		Significance of results (credibility, relevance, and practicalities)	☒	☑			NICE: "often not possible with subgroups" SMC: "From a clinical point of view, important considerations are whether the subgroup has clinical credibility, relevance, and practicality (where it can be easily identified as a group of patients in Scottish practice)"
		If the subgroup is driving the clinical trial results, the indication should be restricted to this group	☑	☑	☑		NICE: "… the drug needs to be cost-effective in the subgroup and not in the overall group. Only then is recommending the drug for a subgroup only appropriate" SMC: "This may particularly be the case for a medicine that looks to have poorer cost-effectiveness for the whole group as we may try to find ways that can maximise the chance of the medicine being accepted at least for some patients" TLV: "if the whole study is driven by a subgroup, then very important to treat this subgroup and to exclude the study as a whole because of evidence demands could be very counter-productive"
		Limited to the marketing authorization and trial indication				☑	
Comparator	Scoping process	By HTA agencies *before* the HTA process (literature review, expert opinion)	☑				
		On the basis of MAH's submission *during* the HTA process (clinical experts queried about choice of comparator)		☑	☑	☑	SMC: "Within the critical appraisal process, SMC will go to a bank of clinical experts with a set of generic questions about the medicine, which tend to elicit responses about comparators, treatments used in current practice, what would be displaced with the new treatment, etc." TLV: "Experts are the most important source of information. As well as guidelines from the Swedish Medical Products Agency about the treatment recommendations for different conditions. The choice of comparator needs to be very specific to Swedish circumstances. Therefore literature reviews does not play"
	Selection	Existing treatment/standard of care to be replaced/routine practice	☑	☑	☑		NICE Guide to Methods of Technology Appraisal TLV: "Criteria for relevant comparator: most cost-effective, treatment most likely to be replaced (e.g. if the patient doesn't get this new drug, what is he/she getting instead)" SMC: Guidance to Manufacturers for Completion of New Product Assessment Form
		Therapeutic technology used at the same stage of the therapeutic strategy				☑	HAS: "therapeutic technologies that you can use at the same stage of the therapeutic strategy"
		Most cost-effective			☑		
Treatment outcomes	Relevant clinical end point, criteria	OS, utility, QALY	☑		☑		NICE only accepts QALYs, and the TLV has a clear preference for QALYs (except when the clinical claim is noninferior efficacy)

continued on next page

Table 2 – *continued*

		NICE	SMC	TLV	HAS	Illustrative quotations/comments
	The end point in the MAH submission is critically appraised (not identified by the committee)		☑			TLV: "TLV is big on QALYs" SMC: "SMC does not identify the end point. The company presents the end point and SMC judges whether it is appropriate or not. This would take place as part of the deliberative process"
	Unmet need based on expert opinion		☑			SMC: "It depends on a range of things such as what was presented in the dossier by the manufacturer, need and unmet need fed from experts"
	End point used for the economic model (e.g., survival plus quality of life for cost-utility analysis)		☑			SMC: "A long-term model would require important information on overall/long-term survival, which is not always possible other than with extrapolation from short-term trials"
	Should reflect the aim of the treatment				☑	HAS: "if the treatment is symptomatic, they will consider the symptoms …"
	Should reflect the short-term consequence of the disease				☑	HAS: "If the disease is leading patients to die shortly, survival should be chosen"
Surrogate end points, acceptability criteria	Validated for life expectancy (= hard end point)	☑	☑	☑	☑	NICE: "If they are not validated against the outcome of interest (qol or life expectancy), they are probably not going to be taken into account" HAS: "If surrogate is validated as predictive for the change of a more hard end point, then it will be accepted"
	Validated for HRQOL (= soft or subjective end point)	☑	☑	☑		SMC: "acceptability is greater where the committee can see that there is an established link between the surrogate outcome measure and the final outcome of interest"
	Clinically relevant			☑		TLV: "Surrogate end points must be clinically relevant. How do they relate to qol and life expectancy"
	Certainty of the validation	☑	☑			NICE: "they will look at the certainty or uncertainty of that validation"
	Nonvalidated	☒		☑	☑	NICE: "They have to be validated" TLV: "If not validated, a surrogate may have to be accepted if it is an important new treatment (and depending on the consequences)" HAS: "If it is not validated, they would not accept surrogates"
	Surrogates for orphan drugs more acceptable	☒	☑	☒	☑	NICE: "we don't differentiate orphan drugs" SMC: "the committee does have more latitude to accept greater uncertainty (through the modifier) and this can lead to a greater acceptance of a surrogate outcome" TLV: "Surrogates are not necessarily more accepted for orphan drugs" HAS: "if there is no other possibility, intermediate end points are accepted"
	Situation of the disease		☑	☒	☑	HAS: "HAS adapt their assessment to the situation. If a disease that has 25–30 patients and the trial has included the same amount of patients in a world-wide situation, they will accept a non-comparative study, with a surrogate end point, etc. They will consider whether they have tried to reach the highest level of evidence they could"
OS and PFS, requirements and preferences	Preference for OS	☑	☑	☑	☑	TLV: "TLV has a preference for OS, but it is hardly the case that that information is available"

continued on next page

	Even if OS is a secondary trial end point	☑				NICE: "It doesn't matter if it is a primary or secondary end point (like utility), NICE will always prefer OS"
	When QALYs depend on life extension		☑			SMC: "OS is preferred where QALY gained depends on life extension"
	When patients may die shortly				☑	HAS: "PFS cannot replace survival in a situation where the patient would die shortly"
PFS, acceptability criteria	PFS validated for OS		☑		☑	HAS: "There is some literature showing that in some kinds of cancer, PFS has shown to be a surrogate of OS and in those cases they would be accepted"
	PFS validated for HRQOL		☑			
	PFS better predictor of (validated) HRQOL than OS		☑			SMC: "For some analyses, PFS is a reasonable outcome to use because it is likely that the main benefits of treatment will be in terms of quality of life rather than in any degree of life extension"
	PFS may be a better predictor of patients' needs (if OS same for two alternatives, area between the curbs may be a value to patients)			☑		TLV: "there might be cases when PFS is at least as interesting and as relevant to patients and clinicians as OS"
HRQOL, requirements and preferences	Required in submission	☑				NICE: Utility measures are needed for the cost-utility model (NICE requirement)
	Preferred if claim is superior efficacy (with a price premium), or if a cost-utility model used (hard end point)		☑	☑		TLV: "We need to have some knowledge of QOL. Or need to make an assumption. Rare are the cases when it not accounted for (e.g. CEA), apart from CMA"
	Preference for generic utility measures (e.g., EQ-5D)	☑	☑			SMC: "SMC has a preference (rather than a requirement) for utility estimates from a validated generic utility instrument such as the EQ 5D"
	Collected within the clinical trial	☑	☑	☑		SMC: "Where utility assessment has taken part within the key clinical studies, we would have a preference for the company using this data in their economic analysis, unless there was a good reason to expect that the data were not appropriate"
	Secondary to assessment (soft end point)				☑	HAS: "HAS will first look at results on the hard end point, and second will look at qol to see how the life is for the patient"
Nongeneric utility measures, acceptability criteria	Validated mapping techniques	☑	☑	☑	☑	SMC: "SMC can accept other sources of utility values, for example, via use of validated mapping techniques or use of values from the literature or registries" TLV: "It is very important to have validated mapping techniques"
	Values from the literature or registries	☑	☑		☑	
	Values from other disease areas	☒	☑	☑	☑	*NICE: "only under exceptional circumstances"
	Values from expert opinions	☒	☑	☑	☑	*NICE: "only under exceptional circumstances" SMC: "The use of expert opinion as a source of utility values would likely be perceived as the most uncertain source of utility values, but has been used in some submissions for some health state valuations" TLV: "Expert opinion can be done for the QALY through the delphi panel (but not when clinical claim is superior efficacy)…It can be used to estimate QALY

continued on next page

Table 2 – *continued*

		NICE	SMC	TLV	HAS	Illustrative quotations/comments
						gains in terms of simplified administration, or parameters that are softer"
Safety assessment, requirements, and preferences	Not assessed per se given it has already been done for marketing authorization	☑	☑	☑		
	Accounted for if it has impact on QALY/utility gains	☑	☑			
	Considered if adequately not captured in the utility values	☑	☑			
	Assessed in the same way as efficacy data				☑	
	Life-threatening diseases, more likely to accept uncertain efficacy if the risk of adverse events is low	☑		☑	☑	NICE: "probably" TLV: "if the risk of severe adverse events from the treatment is low, and that patients can only get better even from taking the treatment, even if we don't know for sure, TLV would allow them to take the chance by paying for this drug. This is also considering that EMA had already looked into safety"

ASMR, relative improvement in clinical benefit ("Amélioration du Service Médical Rendu"); EMA, European Medicine Agency; EQ-5D, EuroQol five-dimensional questionnaire; HAS, Haute Autorité de Santé; HRQOL, health-related quality of life; HTA, health technology assessment; MAH, marketing authorization holder; NICE, National Institute for Health and Care Excellence; OS, overall survival; PFS, progression-free survival; QALY, quality-adjusted life-year; SMC, Scottish Medicines Consortium; SPC, XXX; TLV, Dental and Pharmaceutical Benefits Board.

data such that comparisons become inappropriate (HAS), or poor data quality.

Study population and subgroup data

In general, prespecified subgroup data are preferred by all bodies. Post hoc subgroup data may be accepted if the subgroup corresponds to the licensed or trial indication (HAS) or depending on its relative size and potential significance (SMC). Nevertheless, extrapolating treatment effects from subgroup data to a wider population would not be accepted by HAS. For the other agencies, preference would be given to the subgroup driving the results.

Comparators

Subtle differences were seen in the selection criteria for comparators. NICE uses a scoping process before appraisal to identify the appropriate comparator. The other countries consider the comparators put forward by the MAH. Judgment about their appropriateness is generally based on clinical expertise and local clinical guidance.

Treatment Outcomes and Safety

The choice of relevant clinical end point varies across countries. The SMC judges the appropriateness of the end points provided in the MAH's submission. For HAS, the choice depends on the situation of the disease (e.g., short-term consequences) and the aim of the drug (e.g., symptoms for a symptomatic treatment). NICE and TLV prefer survival data, which would then feed into the quality-adjusted life-year estimate.

The acceptability of surrogate end points depends on their validation against an end point (hard or soft), although HAS insists on a hard end point in such cases. A nonvalidated end point would probably not be accounted for by NICE, whereas it may be accepted under certain circumstances by the others. Surrogate end points for orphan drugs are more likely to be accepted by HAS if no other option is available and by the SMC if it fulfills one of its defined modifiers. NICE and TLV are not more likely to accept surrogate end points for orphan drugs.

There were different levels of acceptability for the surrogate end point progression-free survival. NICE always prefers overall survival to progression-free survival even if it is the trial's secondary end point. Progression-free survival is accepted by the SMC when there is an established link with life extension or the main benefit is improved health-related quality of life (HRQOL). The TLV also prefers overall survival but acknowledges that it is often not available and thus relies on progression-free survival, which is considered potentially closer to patients' needs. For HAS, progression-free survival would not replace overall survival in a situation in which life expectancy is very short unless it were a validated surrogate of overall survival.

HRQOL data were considered either as a hard end point (NICE, TLV, and SMC) or as a soft end point (HAS). The TLV recognized the challenges in collecting HRQOL data for rarer conditions, but the SMC did not consider these challenges to be specific to orphan drugs.

Safety is not part of the assessment for NICE, SMC, and TLV because it is already assessed as part of the drug's marketing authorization. Nonetheless, safety may be considered by the SMC and NICE to the extent it affects quality-adjusted life-year gains or is not adequately captured by utility, survival, and cost estimates. HAS, however, assesses safety along with efficacy. The agencies also agreed that safety may modulate the assessment of efficacy (e.g., if efficacy is the same and safety is worse).

Table 3 – Information about innovation, unmet need, and severity.

HTA body	Innovation	Unmet need	Severity
NICE	Defined by whether the treatment benefits patients, determined during the deliberative process For example, delaying chemotherapy, first oral treatment replacing intravenous administration; counterexamples: new class of drugs, new mechanism of action (without visible benefits to patients)	Defined by the consequence of the decision, determined during the deliberative process in which NICE is willing to accept a higher ICER (up to £30,000/QALY) for the conditions with a high unmet need For example, effect on quality of life of patients without treatment	Defined by the consequence of the decision, determined as part of the deliberative process in which NICE is willing to accept a higher ICER (up to £30,000/QALY) for the more severe diseases. For example, effect on quality of life of patients without treatment
SMC	Intrinsic to the decision, likely captured differently; anything providing benefits to patient, captured by the ICER or accounted for during the deliberative process For example, a first in class could fulfill an unmet need, new mode of action or administration benefits, advantages in terms of service delivery, reduced severe adverse events, step change in patient management	For orphan drugs through the modifiers, "lack of available treatments of proven efficacy," determined as part of the deliberative process and from clinical experts "No treatment" would be prioritized over "few treatments." If there were "few treatments" with intolerable side effects, it would be considered an unmet need	No definition, may be accounted for intrinsically during the deliberative process
TLV	Benefits to patients, captured by the ICER or as part of the deliberative process For example, improved administration form benefits patients and reduces costs	Defined by the consequence of the decision, determined as part of the deliberative process Disease severity and unmet need are considered to be related: the greater the severity, the greater the unmet need	
HAS	Captured within the ASMR; a drug with an ASMR I, II, or III would be considered as innovative. Prices would be set at European levels and would not be negotiated with the economic committee (CEPS)	Captured within the SMR; place in the therapeutic strategy: if no other options at the same stage of the disease, on the basis of the analysis of comparators and the description of therapeutic strategy (how the disease is treated, where the drug would fit, and what are the current existing alternatives) For example, a real unmet medical need would be recognized when there are no other treatment options	Captured within the SMR; different categories of severity defined: severe, life-threatening, short life expectancy, affects quality of life, not so severe

ASMR, relative improvement in clinical benefit ("Amélioration du Service Médical Rendu"); CEPS, economic committee ("Comité Economique des Produits de Santé"); HAS, Haute Autorité de Santé; HTA, health technology assessment; ICER, incremental cost-effectiveness ratio; NICE, National Institute for Health and Care Excellence; QALY, quality-adjusted life-year; SMC, Scottish Medicines Consortium; SMR, clinical benefit ("Service Médical Rendu"); TLV, Dental and Pharmaceutical Benefits Board.

Additional Qualitative Criteria Considered

The interviews provided additional insights into the relevant evidence and weight of each of the criteria innovation, unmet need, and disease severity (Table 3). NICE explicitly accounts for innovation, defined as a step change for patients rather than as a new class of drug or mechanism of action, whereas the TLV and SMC do not have specific criteria. HAS gives innovative drugs higher ASMR ratings (I, II, III), which result in European pricing rather than a lower negotiated price. Likewise, innovative hospital drugs are covered over and above applicable diagnosis related group (DRG) tariffs.

For NICE, severity is accounted for in drugs with high ICERs by looking at how a patient's HRQOL is affected without treatment rather than improved survival, which is considered to be captured by the model with its baseline severity. The TLV also explicitly accounts for severity in accepting higher ICERs. Disease severity corresponds to a greater unmet medical need, although no explicit weighing or definition of severity currently exists. The SMC does not explicitly account for severity, although committee members may intrinsically capture this element. For HAS, severity is incorporated into its Service Médical Rendu (SMR) ratings that drive coverage levels. Predefined categories of severity have been defined: severe, life-threatening, short life expectancy, affects quality of life, not so severe.

NICE and TLV consider disease severity and unmet need together by focusing on the consequence of refusal to cover the drug. NICE also accounts for the drug's place in the therapeutic strategy given the medical need and would recognize an unmet need when no other treatment options are available. The SMC

assesses unmet need on a case-by-case basis, drawing on clinical expertise to understand current treatment options and how the new treatment would fit into clinical practice. Unmet need would be recognized through the application of a decision modifier ("lack of available treatments of proven efficacy"), which is strictly applied when there is no proven treatment available for a particular indication. Thus, a situation in which there is no treatment at all would likely have priority over one for which few treatments exist. The existence of intolerable side effects would also give rise to a recognition of unmet need. HAS considers unmet need in the context of assessing the place of the treatment in the therapeutic strategy, which includes identification of comparators. If no other options were available, this would be considered a great unmet need.

Discussion

This study aimed to elicit the views of HTA bodies about their approaches to valuing orphan drugs. HTA bodies agreed that the evidence for orphan drugs is of lower quality than that for drugs for more prevalent conditions. Despite the broad agreement seen in evidentiary requirements or preferences, subtle differences were identified with respect to the circumstances under which this imperfect or incomplete evidence may be considered acceptable, which may influence HTA outcomes and explain differences across countries.

Despite the known limitations in generating robust evidence for orphan drugs [4,9], formal evidentiary requirements are similar for orphan and nonorphan drugs, with the exception of HAS, which accounts for the situation of the disease, when a small trial population and noncomparative trial would be considered acceptable if the number of patients living with the disease was very low. The TLV and HAS have higher evidentiary requirements for superior efficacy, which has implications for orphan drugs when the lack of treatment options increases the likelihood that a new treatment would be superior [6]. Demonstrating survival or other clinically relevant patient benefits requires well-designed large trials [17,18]. For orphan drugs, greater treatment effects would be required from small-scale trials to attain statistical significance [19], unless innovative small-scale trial designs are used [20].

Countries applied similar criteria in assessing appropriate trial duration, which were related to the natural course of the disease and likely duration of the treatment. Challenges, however, exist in defining the appropriate trial duration for orphan drugs, particularly when the natural history of the disease is unknown [4] or when the disease is chronic or has an early onset. Orphan drugs are often characterized by short clinical testing phases [6], which may not be sufficiently long to capture their benefits in clinical practice, particularly for lifelong conditions.

Registries are particularly useful for rare diseases by collecting information about the patient experience and natural history of the disease, thereby improving the quality and reliability of this evidence [21]. The use of registries, however, remains limited because of their time-consuming nature and challenges in analyzing historical evidence of a product's effectiveness [22]. This was confirmed by the interviewees.

Regarding subgroups, the HTA bodies generally preferred prespecified subgroups and imposed restrictions when cost-effectiveness was driven by a subgroup of patients. The situation was different for HAS, which requires subgroups to be similar to those considered for marketing authorization or included in the trial. Orphan drugs often treat rare diseases of genetic origin that affect children. Of the 81 orphan drugs receiving marketing authorization from the European Commission after scientific review and recommendation by the European Medicine Agency since 2000, half are authorized for a subgroup of children and another 34 are under investigation in children [23]. In addition, 30% to 40% of orphan drugs treat different cancers, which are characterized by an increasing body of research on predictive biomarkers to assess treatment response [24]. Despite this, very few subgroups of different subtypes are included in licensing indications and those that do may fail to reflect clinical practice [25]. A review of 894 randomized controlled trials showed that half reported subgroup analyses, of which 46% were planned in the trial protocols and only 10% matched those reported in the publication [26]. Thus, subgroup data must be assessed with caution. Given the frequency of subgroup data in rare diseases because of their heterogeneity, lack of knowledge about existing subtypes [4], and licensing of orphan drugs for adult patients in diseases that commonly affect children [27], new methods are needed to address these issues beyond simply restricting the indication, such as imposing re-assessments, providing coverage contingent upon evidence development, and collecting real-world data through registries.

Issues regarding comparative evidence are more frequent for orphan drugs given that they often rely on single-arm, nonrandomized studies [6] and that expertise about clinical pathways may be lacking [4]. This also implies that there are likely to be greater differences across countries in their definitions of standard care pathways. Therefore, it may be even more challenging to produce comparative trials with the appropriate comparators for a particular clinical context.

Surrogate end points are more common for orphan drugs than for nonorphan drugs [6,28]. Differences in acceptability of end points for validation (hard vs. soft) may have implications for orphan drugs, given questions about their clinical relevance (e.g., improvement in walking or platelet response) [29] or difficulties in establishing their validation [19]. Evidence suggests that surrogate end points tend to overestimate treatment effects, which may be minimized by quantifying their magnitude and validating them with relevant patient outcomes [30]. This issue also underscores the need to ensure that what is being measured for HTA is responsive to patient needs, preferences, and values through continuous involvement of patients in the drug development process [31]. Patient input could help determine, for example, whether overall survival or progression-free survival responds better to patient needs in a particular disease setting.

Some differences were seen in the way the qualitative criteria were defined by HTA bodies, and interviewees acknowledged inconsistency in their application within the deliberative process. This may originate either from the different approaches adopted by each agency, which were identified in this study, or from individual appraisal committee members' judgments on the basis of their experience or on what they believe society would prefer [10]. Despite agreement in defining innovation in terms of benefits to patients, differences were seen in accounting for innovation: explicitly by NICE and HAS through the ASMR and by TLV and SMC (or implicitly as a value judgment of committee members) through the ICER. Definition of unmet need also differed among the HTA bodies, with NICE and TLV defining it as a consequence of the decision, and the SMC and HAS defining it as the lack of available treatment options. The first accounts for disease severity and impact on patients without treatment, whereas the second accounts for the availability of alternative treatments without differentiating for severity. Although severity is captured together with unmet need for NICE and TLV, it is not explicitly accounted for by the SMC and HAS. HAS does account for different severity criteria, but their influence on the assessments has been shown to be minimal (e.g., severe disease identified in 50% of drugs with insufficient SMR) [32].

These differences suggest the need for better defined qualitative criteria, which would also improve consistency in their

application. Innovation could be defined as those aspects that benefit patients but are not captured by the ICER or the ASMR, including considerations around managing and living with a disease. The key issue surrounding unmet need is whether it should capture severity and prioritize the most severe conditions or whether it is a way to ensure that patients have treatment options at each disease stage. Prioritizing the most severe conditions would give less weight to certain "less severe" or non–life-threatening problems from living with the disease or taking the treatment (e.g., pain, adverse effects, and reduced mobility).

Evidentiary requirements for cost-effectiveness (NICE, SMC, and TLV) differ from clinical benefit assessments (HAS). For example, differences were seen in whether HRQOL data were considered a hard or soft end point. As a soft end point, lack of improvement in HRQOL data may have greater implications than if captured explicitly within an economic model. Furthermore, for cost-utility modeling, generic utility data are generally preferred despite not always being the most appropriate way to capture quality of life [33]. This may have implications for rare diseases, which are often chronic, severe, and disabling, affecting HRQOL and other aspects of life, such as hopelessness linked to illness chronicity or the search for normalcy and social recognition as part of a community [34].

The structure of the qualitative research review guidelines (RATS) was followed to ensure the quality and clear dissemination of this research [35]. Nonetheless, the study is not without limitations. First, the interview questions were derived from the analysis of 10 orphan drugs. Although this sample may not be representative of all issues surrounding orphan drugs, the scenarios encountered were repeated, suggesting that the most common types of issues encountered were captured. The main advantage of focusing the interview questions on scenarios is that it allowed comparison with what was seen in practice. Second, varying levels of detail may have been captured during the interviews. To address this, a second round of questions with tables summarizing responses to the initial interview was sent to all interviewees to ensure comparability and reliability of the research. Third, the differentiation between how these findings apply to orphan and nonorphan drugs, which undergo the same assessment process, was at times unclear. We, nevertheless, were able to identify certain issues specific to orphan drugs and to explore how the process could be adapted to overcome some of these unique challenges.

Conclusions

Orphan drugs, which generally are subject to the same processes as drugs for more prevalent conditions, are assessed in a context of lower quality evidence, and this study contributed to understanding how HTA bodies address these challenges. Although agreement was seen regarding evidentiary requirements and preferences, differences were apparent in how this imperfect or incomplete evidence was considered, which may explain conflicting recommendations. This study further identified systemic features that are not well adapted to assessment of orphan drugs, which may need to be reconsidered to ensure that their value is appropriately captured when used to inform reimbursement decisions. This is all the more compelling in a pharmaceutical environment that is shifting toward more niche and targeted therapies in which HTA bodies will increasingly face such issues.

Acknowledgments

We thank the European Commission for supporting this project. A special thanks also goes to the interviewees for their time and valuable contributions to this article.

Source of financial support: This research was funded by the European Commission's 7th Framework Programme and undertaken under the auspices of Advance-HTA (grant no. 305983). The results presented here reflect the authors' views and not those of the European Commission. The European Commission is not liable for any use of the information communicated.

Supplemental Materials

Supplemental material accompanying this article can be found in the online version as a hyperlink at http://dx.doi.org/10.1016/j.jval.2017.03.005 or, if a hard copy of article, at www.valueinhealthjournal.com/issues (select volume, issue, and article).

REFERENCES

[1] Hurley J. An overview of the normative economics of the health care sector. In: Culyer A, Newhouse J, eds. Handbook of Health Economics. Amsterdam, The Netherlands: Elsevier, 2000.
[2] Caro JJ. Pursuing efficiency: A dead end for HTA? Value Health 2009;12:S49.
[3] 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.
[4] Vickers PJ. Challenges and opportunities in the treatment of rare diseases. Drug Discov World 2013;14:9–16.
[5] Bell S, Smith C. A comparison of interventional clinical trials in rare versus non-rare diseases: an analysis of ClinicalTrials.gov. Orphanet J Rare Dis2014;9:170.
[6] Kesselheim AS, Myers JA, Avorn J. Characteristics of clinical trials to support approval of orphan vs nonorphan drugs for cancer. J Am Med Assoc 2011;305:2320–6.
[7] Drummond MF, Wilson DA, Kanavos P, et al. Assessing the economic challenges posed by orphan drugs. Int J Technol Assess Health Care 2007;23:36–42.
[8] Nicod E, Kanavos P. Developing an evidence-based methodological framework to systematically compare HTA coverage decisions: a mixed methods study. Health Policy 2016;120:35–45.
[9] 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. EU J Health Econ (published online ahead of print August 18, 2016). http://dx.doi.org/10.1007/s10198-016-0823-0.
[10] Nicod E, Kanavos P. Scientific and social value judgments for orphan drugs in health technology assessment. Int J Technol Assess Health Care 2016;32:218–32.
[11] Advance-HTA. Rethinking the future of health technology assessment. Available from: www.advance-hta.eu. [Accessed 14 June 2015].
[12] Scottish Medicines Consortium. SMC modifiers used in appraising new medicines. Available from: https://www.scottishmedicines.org.uk/About_SMC/Policy_statements/SMC_Modifiers_used_in_Appraising_New_Medicines. [Accessed 15 February 2017].
[13] Gale NK, Heath G, Cameron E, et al. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Med Res Methodol 2013;13:17.
[14] Rawlins MD. Evidence, values, and decision making. Int J Technol Assess Health Care 2014;30:233–8.
[15] Rawlins MD, Culyer AJ. National Institute for Clinical Excellence and its value judgments. Br Med J 2004;329:224–7.
[16] Claxton K. Exploring uncertainty in cost-effectiveness analysis. Pharmacoeconomics 2008;26:781–98.
[17] Korn EL, Freidlin B, Abrams JS, et al. Design issues in randomized phase II/III trials. J Clin Oncol 2012;30:667–71.
[18] Wieand HS. Randomized phase II trials: What does randomization gain? J Clin Oncol 2005;23:1794–5.
[19] Boudes PF. Clinical studies in lysosomal storage diseases: past, present, and future. Rare Dis 2013;1:e26690.
[20] Gagne JJ, Thompson L, O'Keefe K, et al. Innovative research methods for studying treatments for rare diseases: methodological review. Br Med J 2014;349:g6802.
[21] Howie L, Hirsch B, Locklear T, et al. Assessing the value of patient-generated data to comparative effectiveness research. Health Affair 2014;33:1220–8.
[22] Haffner ME. Designing clinical trials to study rare disease treatment. Drug Inf J 1998;32:957–60.

[23] Kreeftmeijer-Vegter AR, de Boer A, van der Vlugt-Meijer RH, et al. The influence of the European paediatric regulation on marketing authorisation of orphan drugs for children. Orphanet J Rare Dis 2014;9:120.
[24] Kelly C, Kelly CM. Understanding molecular pathways and the identification of subgroups of patients with specific biomarker predictors of drug response related to orphan drugs. Expert Opin Orphan Drugs 2013;1:67–76.
[25] Malottki K, Biswas M, Deeks JJ, et al. Stratified medicine in European Medicines Agency licensing: a systematic review of predictive biomarkers. Br Med J Open 2014;4:e004188.
[26] Kasenda B, Schandelmaier S, Sun X, et al. Subgroup analyses in randomised controlled trials: cohort study on trial protocols and journal publications. Br Med J 2014;349:g4539.
[27] Aronson JK. Rare diseases and orphan drugs. Br J Clin Pharmacol 2006;61:243–5.
[28] Joppi R, Bertele V, Garattini S. Orphan drugs, orphan diseases: the first decade of orphan drug legislation in the EU. Eur J Clin Pharmacol 2013;69:1009–24.
[29] Joppi R, Bertele V, Garattini S. Orphan drug development is not taking off. Br J Clin Pharmacol 2009;67:494–502.
[30] Ciani O, Buyse M, Garside R, et al. Comparison of treatment effect sizes associated with surrogate and final patient relevant outcomes in randomised controlled trials: meta-epidemiological study. Br Med J 2013;346:f457.
[31] Facey K, Hansen HP. The imperative for patient-centred research to develop better quality services in rare diseases. Patient 2015;8:1–3.
[32] Polton D. Réforme sur la réforme des modalités d'évaluation des médicaments. Ministère des Affaires sociales et de la santé 2015. http://social-sante.gouv.fr/IMG/pdf/rapport_polton_-_evaluation_medicaments.pdf. [Accessed 15 June 2016].
[33] Tordrup D, Mossman J, Kanavos P. Responsiveness of the EQ-5D to clinical change: Is the patient experience adequately represented? Int J Technol Assess Health Care 2014;30:10–9.
[34] Caputo A. Exploring quality of life in Italian patients with rare disease: a computer-aided content analysis of illness stories. Psychol Health Med 2014;19:211–21.
[35] Clark J. How to peer review a qualitative manuscript. In: Godlee F, Jefferson T, eds., Peer Review in Health Sciences (2nd ed.). London: BMJ Books, 2003.