


ScienceDirect

Contents lists available at **sciencedirect.com**
Journal homepage: **www.elsevier.com/locate/jval**

Themed Section: HTA Around the World—Influences of Culture, Values, and Institutions

Differences in Health Technology Assessment Recommendations Among European Jurisdictions: The Role of Practice Variations

Rick A. Vreman, PharmD,[1,2] Aukje K. Mantel-Teeuwisse, PhD,[2] Anke M. Hövels, PhD,[2] Hubert G.M. Leufkens, PhD,[2] Wim G. Goettsch, PhD[1,2,*]

[1]*Division of Pharmacoepidemiology and Clinical Pharmacology, Utrecht Institute for Pharmaceutical Sciences, Utrecht University, Utrecht, The Netherlands;* [2]*The National Healthcare Institute, Diemen, The Netherlands.*

ABSTRACT

Background: Health technology assessment (HTA) plays an important role in reimbursement decision-making in many countries, but recommendations vary widely throughout jurisdictions, even for the same drug. This variation may be due to differences in the weighing of evidence or differences in the processes or procedures, which are known as HTA practices.

Objective: To provide insight into the effects of differences in practices on interpretation of intercountry differences in HTA recommendations for conditionally approved drugs.

Methods: HTA recommendations for conditionally approved drugs (N = 27) up until June 2017 from England/Wales, France, Germany, the Netherlands, and Scotland were included. Recommendations and practice characteristics were extracted from these five jurisdictions and this data was validated. The effect of nonsubmissions, resubmissions, and reassessments; cost-effectiveness assessments; and price negotiations on changes in the percentage of negative recommendations and the interpretation of intercountry differences in HTA outcomes were analyzed using Fisher exact tests.

Results: The inclusion of cost-effectiveness assessments led to significant increases in the proportion of negative recommendations in England/Wales (from 4% to 50%, $P<.01$) and Scotland (from 21% to 71%, $P<.01$). The subsequent inclusion of price negotiations led to significant reductions in the proportion of negative recommendations in England/Wales (from 50% to 14%, $P<.01$), France (from 31% to 3%, $P=.012$), and Germany (from 34% to 0%, $P<.01$). Results indicated that the inclusion of nonsubmissions and resubmissions might affect Scottish negative HTA recommendations (from 7% to 21%), but this effect was not significant. No significant effects were observed in the Netherlands, possibly owing to sample size.

Conclusion: Variations in HTA practices between international jurisdictions can have a substantial and significant impact on conclusions about recommendations by HTA bodies, as exemplified in this cohort of conditionally approved products. Studies comparing international HTA recommendations should carefully consider possible practice variations between jurisdictions.

Keywords: conditional marketing authorization, cost-effectiveness assessment, differences, evidence, health technology assessment, practice, procedure, process, regulation, reimbursement, relative-effectiveness assessment.

VALUE HEALTH. 2020; 23(1):10–16

Introduction

Recommendations on the reimbursement of health technologies are provided by health technology assessment (HTA) bodies. Because reimbursement is a regional or national process, many different HTA organizations exist in Europe. Although most of these HTA organizations collaborate in the European Network for Health Technology Assessment (EUnetHTA) on joint relative effectiveness assessments (REAs),[1] each HTA organization ultimately provides recommendations or decisions for constituents within their own jurisdiction and based on their regional or national preferences. Consequently, differences in HTA recommendations have been the subject of extensive research.[2–19]

Studies qualitatively assessing recommendations for groups of drugs in several HTA jurisdictions report that a series of sources contribute to observed differences. These sources can be roughly divided into two categories. First, there are inconsistencies in practices, involving the ability to implement patient access

* Address correspondence to: Wim G. Goettsch, PhD, Utrecht Institute for Pharmaceutical Sciences, Utrecht University, Universiteitsweg 99, 3584 CG, Utrecht, the Netherlands. Email: w.g.goettsch@uu.nl

https://doi.org/10.1016/j.jval.2019.07.017

schemes, the ability to file resubmissions/reassessments and nonsubmissions, and the HTA approach applied (eg, inclusion or omission of cost-effectiveness).[7,9] These inconsistencies can be classified as practice differences because they relate to the different procedures, processes, laws, tactics, and other factors applicable in different jurisdictions. Second, even when jurisdictions apply the same practice (eg, a cost-effectiveness assessment [CEA]), there may still be differences in HTA outcomes, stemming from differences in the inclusion and evaluation of evidence, the impact of evidence on the final recommendation, and the integral leveraging of different types of evidence.[2-6,8] These can be classified as value assessment differences. Within HTA practices, a nonsubmission means that an HTA body provides a recommendation on reimbursement, even though no submission has been provided by the relevant company. Products can be assessed more than once, either because the manufacturer files a resubmission or because an HTA body decides to perform a reassessment. Some studies include resubmissions and reassessments,[5] whereas others do not or only do so in additional analyses.[9,10] Interpreting these studies should be done with caution because results may not be comparable. Differences in the timing of recommendations have also been established among and within countries.[11] Studies including multivariable analyses have tried to specify the relative importance of these factors on HTA recommendations. The importance of factors depends, however, on the jurisdiction(s) assessed and the variables added to the model. The inclusion of explanatory variables in univariate and multivariable models varies widely, and results are simultaneously variable and sometimes contradictory.[12-17] The inclusion of practice-related factors is rare.[18]

Importantly, overall HTA recommendations are the result of very different practices in international HTA jurisdictions. These practices have been shaped by societal and cultural values and by the development of the role of HTA in healthcare decision making within individual jurisdictions. The existence of differences in HTA practices is a logical consequence of divergent values. A study assessing the importance of criteria to decision makers in multiple countries has quantified some of these differences, showing that criteria (eg, cost-effectiveness) may be valued extremely high in some jurisdictions but deemed irrelevant in others.[11] The HTA practices resulting from these values thus diverge. For example, in England/Wales and Scotland, most overall recommendations include CEAs and price negotiations. In the Netherlands, some HTA recommendations include cost-effectiveness evaluations, but none include pricing negotiations. In France, cost-effectiveness evaluations are relatively rare and in Germany they are absent. Nevertheless, whether results are positive, negative, or restricted in these last two jurisdictions depends on how the added benefit assessment (relative effectiveness) outcomes are classified. For example, when comparing French or German recommendations to English/Welsh recommendations, it makes sense to qualify the French and German categories that result in lower pricing benchmarks as (economically) restricted. Alternatively, if one is interested in the results of REAs independent of concomitant pricing restrictions, one would classify categories that define minor-major added benefit as positive.

In conclusion, differences in HTA recommendations arise because of a divergence in how evidence informs evaluation (value assessment differences) and owing to disparities in how regional or national values have shaped HTA practices (practice differences).[7,19] Although value assessment differences have been substantially categorized and quantified, the quantification of the effects of practice differences is lacking. Currently, it is unclear to what extent practice differences may affect overall conclusions on international comparisons of HTA recommendations.

In previous research, substantial differences were shown between HTA jurisdictions in recommendations for products approved by the European Medicines Agency (EMA) through conditional marketing authorization (CMA).[10] The present study aims to provide insight into the contribution of differences in practices on the observed divergence. More specifically, this study aims to clarify whether appropriately considering differences in HTA practices can lead to significantly altered conclusions in research when comparing international HTA recommendations, using conditionally approved products as an example.

Methods

Selection of Medicines and Jurisdictions

All drugs approved under CMA between January 2006 and June 2016 were included. The year 2006 represents the year of the implementation of the CMA scheme. Published HTA recommendations were included up until June 2017 to permit some time for assessment after approval. Jurisdictions were selected based on four previously published criteria[10]: (1) the HTA jurisdiction had to be linked to an EU jurisdiction, (2) reports had to systematically be publicly available, (3) the HTA body involved had to be the primary institute with legal remits within the jurisdiction, and (4) the report had to be in a language understood by the assessors (ie, English, French, German, or Dutch). This led to the inclusion of HTA recommendations from England/Wales (NICE), France (HAS), Germany (IQWIG), Scotland (SMC), and the Netherlands (ZIN).

Data Collection

Data on HTA recommendations for medicines were collected from HTA reports through a standardized data extraction format. HTA recommendations were matched to primary indications approved by the EMA. Extraction was performed by the first author and validated by a second researcher through independent extraction for a random sample of 10% of included drugs. Recommendations for drugs with multiple indications were included separately, as were recommendations that were explicitly split by HTA bodies because multiple comparators existed. Data collected included basic drug characteristics (generic and brand name, orphan status, date of regulatory approval, indication, and therapeutic category), HTA practice characteristics (date of recommendation/publication of recommendation, submission/recommendation type, number and outcome of submissions/resubmissions, and inclusion or exclusion of price negotiations and CEAs in the recommendation), and assessment outcomes (overall recommendation, REA, CEA, reasons for not recommending, and descriptions of restrictions). For the purpose of analysis, these characteristics and outcomes were coded on an individual drug level. The data extraction was part of a larger study on HTA recommendations.[10] Interrater agreement was 90.7%, with the unweighted Kappa being 0.867, indicating excellent agreement.[20,21] Discrepancies were resolved through discussion.

Analysis

In line with the objectives of this study, the analysis is designed to show the effects of differences in practices on overall HTA recommendations. As previously mentioned, many differences in HTA practices exist. Nevertheless, previous research has highlighted three practices that represent a large part of the variation in the application of HTA between jurisdictions: (1) the possibility of resubmissions, reassessments, and nonsubmissions (together called reassessments in this analysis), (2) the ability to include a

CEA, and (3) the ability to include price negotiations in the final HTA recommendation.[7,9] Although many other variations in practices exist, the aim of this study was to show the potential impact of practice differences on international comparisons of HTA outcomes. It was expected that these three major effects would be sufficient to establish this effect. The included jurisdictions vary greatly in their application of HTA, as outlined by previous research, which provided classification systems for HTA institutions.[22]

The first practice variation that is included is the possibility for resubmissions, reassessments, and nonsubmissions. All included jurisdictions have processes in place that facilitate reassessments based on different triggers. Triggers may comprise a change to the marketing authorization, a specific request by stakeholders, or a planned reassessment, among other possibilities. Additionally, companies can provide a resubmission of a reimbursement dossier when they feel new data warrants a reassessment. Even if a company does not submit a dossier, some jurisdictions (ie, England/Wales and Scotland) may issue so-called nonsubmissions, which provide an HTA recommendation (usually a negative one) regardless of the lack of a dossier. When comparing jurisdictions, it is crucial to decide whether to include the outcomes of reassessments and nonsubmissions because HTA recommendations may differ between the primary and subsequent assessments.

The second practice refers to the inclusion of CEAs. Some jurisdictions include cost-effectiveness as a standard criterium within their HTA process (ie, England/Wales and Scotland) whereas others do not perform CEAs (Germany) or only perform them sometimes, based on a risk-assessment (France and the Netherlands). The inclusion of CEAs may lead to very different overall HTA recommendations as opposed to only including REAs.

The third included practice is the possibility of including economic restrictions, such as price negotiations, in the assessment process. Some jurisdictions may include price negotiations within the HTA process (ie, England/Wales and Scotland), whereas others never do (the Netherlands). Alternatively, the HTA process itself may lead to economic restrictions because the REA may result in a suggestion for a benchmark price to be used in negotiations (Germany, France, and the Netherlands).

Thus we established four groups of HTA recommendations based on the three established variations in HTA practices. The groups are clarified in Table 1. The first group includes only the first published REA and is called REA. Because REA is a practice performed by all jurisdictions, it can serve as a reference practice. HTA recommendation outcomes were categorized as positive, positive with restrictions (called restricted), or negative, in line with previous research.[4] CEAs were separated from REAs in jurisdictions performing both. Reported REAs in HTA dossiers may entail three (positive, restricted, negative) to six (ranging from less benefit to major benefit) categories and vary per jurisdiction. For group one, we categorized relative-effectiveness outcomes that stated benefit over the comparator as positive, even if that benefit was minor. A lack of benefit or less benefit was qualified as negative. In the second group (called SUB), more submissions were taken into account. Inclusion was broadened by adding both HTA recommendations based on nonsubmissions and the final (last) recommendation after resubmissions and reassessments, which led to changes in the recommendation for some drugs. The third group is called CEA and includes CEAs, the REA and resubmissions/reassessments. It was expected that this would result in more negative HTA recommendations because positive REAs can be followed by a negative CEA. Finally, in the fourth group (called PRI), price negotiations and economic restrictions are included, which could reduce the proportion of negative recommendations. In countries not performing price negotiations within the HTA process, certain REA or CEA outcomes may still lead to economic restrictions, such as product clustering with maximum prices (the Netherlands) or pricing benchmarks (France and Germany). The REA and CEA outcomes that are associated with such economic restrictions are defined as restricted in group 4.

Table 1. Variations in HTA practices included within each defined group of HTA recommendations.

	REA	SUB	CEA	PRI
First relative effectiveness assessment	Yes	Yes	Yes	Yes
Reassessments, resubmissions, or nonsubmissions	No	Yes	Yes	Yes
First and subsequent cost-effectiveness assessments	No	No	Yes	Yes
Economic restrictions or price negotiations	No	No	No	Yes

HTA indicates health technology assessment; REA, abbreviation for practice 1; SUB, abbreviation for practice 2; CEA, abbreviation for practice 3; PRI, abbreviation for practice 4.

The significance of the differences in the proportion of negative recommendations was tested using Fisher exact tests. Only the recommendation outcomes changed between groups 2, 3, and 4 and not the amount of included recommendations. Between groups 1 and 2, the number of assessed products changed owing to the inclusion of nonsubmissions. Thus, significance was tested both by including non-assessed products and by excluding non-assessed products. The comparison of assessed products (disregarding non-assessments) is a common approach in international comparisons of HTA recommendations.[10]

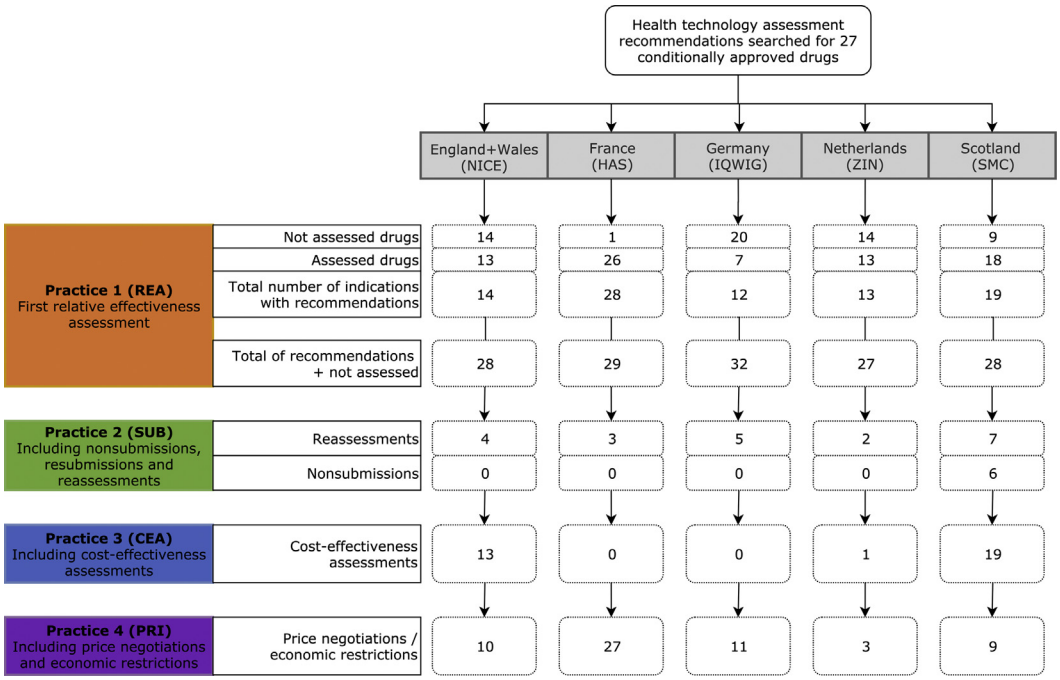
To provide insight into the effects of practice differences on international comparisons of HTA recommendations, the ranking of countries on a scale that reports the proportion of negative recommendations was investigated.

Results

Thirty CMA drugs were approved by the EMA up until June 2016, of which 3 were excluded from this study because they had not been assessed by any HTA organization. In total, 92 HTA recommendations were available on 30 June 2017 for the remaining 27 drugs, spanning 34 indications. Practice differences in the selection of drugs for assessment became apparent through the variation in the number of drugs assessed because primary recommendations (REAs) were present in Germany for only 7 drugs (26%), whereas in France, they were present for 26 (96%). The included numbers of recommendations per practice per jurisdiction are shown in Figure 1. Note that some jurisdictions make separate recommendations for each indication when drugs have multiple indications. Thus the total number of indications with recommendations may exceed the number of drugs assessed.

Figure 2 shows the HTA outcomes for the different practices. Differences between practices 1 and 2 (REA and SUB) within jurisdictions represent the effect of the inclusion of nonsubmissions, resubmissions, and reassessments. Within our group of HTA bodies, nonsubmissions were only possible in Scotland. Thus the biggest impact is noted in this jurisdiction. The proportion of negative recommendations by the SMC changed from 7% to 21% (P=.252). Differences in other jurisdictions were small (Germany) or absent (all others). The inclusion of CEAs (as in practice 3, CEA) did not have an effect on French and German recommendations because no CEAs were available for these jurisdictions in our data

Figure 1. Included HTA recommendations per practice. Some jurisdictions provide multiple recommendations when drugs have several indications, resulting in more recommendations than drugs.



CEA indicates cost-effectiveness assessment; HTA, health technology assessment; REA, relative effectiveness assessment.

set. From the Netherlands, only one CEA was included, which was positive and led to no change. In the other two jurisdictions, the inclusion of CEAs led to more negative recommendations. This effect is significant between practices 2 and 3, not only when excluding non-assessed indications (England/Wales went from 7% to 100% and Scotland from 24% to 80%; for both, $P<.01$) but also when including the non-assessed indications (England/Wales went from 4% to 50% and Scotland from 21% to 71%; for both, $P<.01$). Alternatively, when price negotiations were included as well (as in practice 4, PRI), these resulted in more (economically) restricted positive outcomes. The change in proportion of negative recommendations excluding non-assessed indications was significant in England/Wales (from 100% to 29%, $P<.01$), France (from 32% to 4%, $P=.012$), and Germany (from 92% to 0%, $P<.01$) between practices 3 and 4. When non-assessed products were included, significance remained ($P<.01$, P=.012, and $P<.01$, respectively). The size of the effects of each practice on recommendation outcomes varied widely per jurisdiction. HTA recommendations for all the jurisdictions combined are also depicted in Figure 2, which shows that the large effects present in individual jurisdictions may be ameliorated by a lack of effect in other jurisdictions, resulting in fewer changes in the combined set (eg, between practices 2 and 3). Alternatively, it is possible that an effect shown through an analysis of all the jurisdictions can be completely explained by the effects in one jurisdiction (eg, between practices 1 and 2).

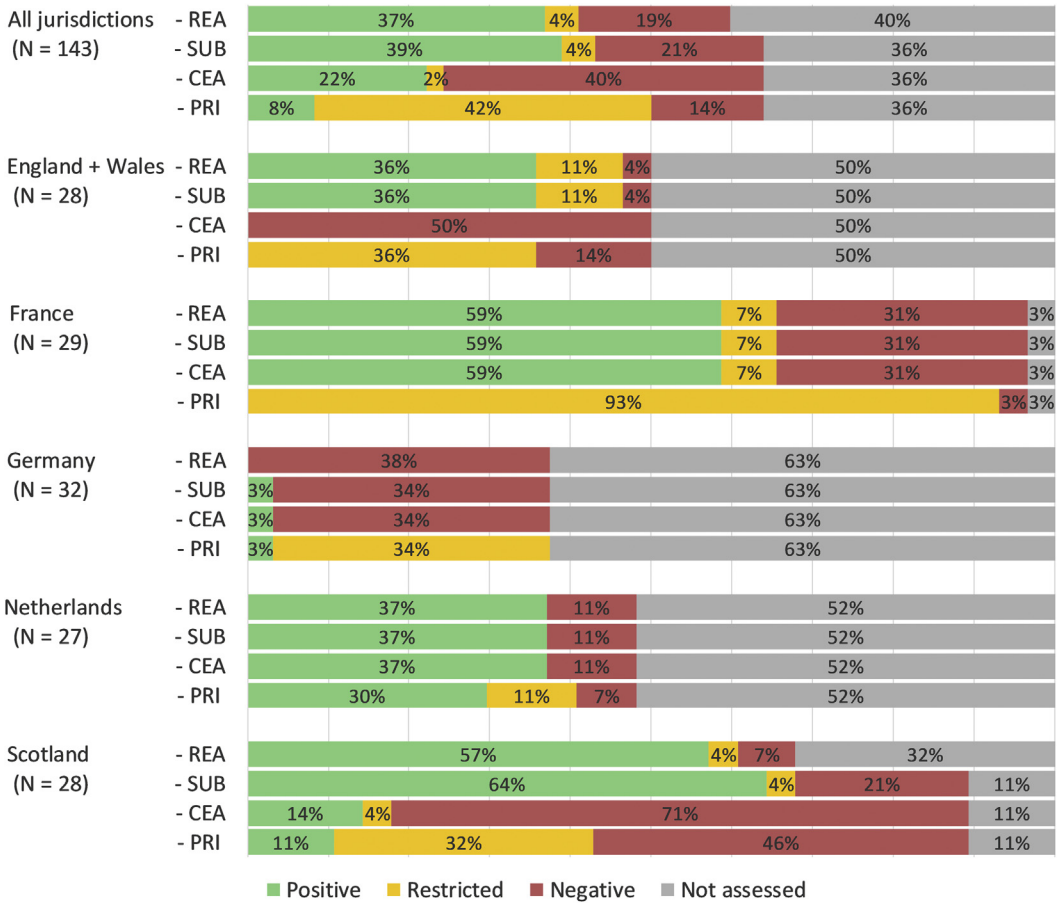
The effects of practice differences on comparisons between jurisdictions is further explored in Figure 3. As Figure 2 already indicates, the proportion of negative recommendations per jurisdiction can change substantially if practice differences are accounted for. Figures 3 and 4 show that rankings of jurisdictions based on the proportions of recommendations that are negative are substantially altered between the practices that we have defined. Including CEAs in a comparison (practices 3 and 4; CEA and PRI) resulted in relatively high proportions of negative recommendations for England/Wales and Scotland, whereas these two jurisdictions had the lowest proportion of negative recommendations when CEAs and price negotiations were excluded (practices 1 and 2; REA and SUB). A comparison between Figures 3 and 4 shows how rankings are affected when non-assessed products are included versus when they are excluded. For Figure 3, it is impossible for proportions of negative recommendations to reach 100% because no recommendations exist for non-assessed products.

Discussion

Our results show that taking HTA practice differences into account can have a substantial impact on conclusions regarding HTA recommendations for conditionally approved products in European jurisdictions. Within individual jurisdictions, HTA recommendations can be significantly altered when distinct practices such as REAs and CEAs are considered separately. Comparing overall HTA recommendations between jurisdictions without accounting for practice differences within those jurisdictions can lead to substantially altered conclusions on the jurisdictions' negativity.

We used conditionally approved products as a case study to clarify the effects of differences in practices. Because no specific HTA processes exist for conditionally approved products, it may be assumed that the practice differences we investigated will also be relevant for other types of products. The extent to which they are relevant, however, may be different. We have considered three relevant practice differences established in previous research. Nevertheless, many more practice variations exist that may have effects, starting with the selection procedure for products to be assessed (see Figures 3 and 4). The relevance and impact of each practice variation on outcomes of comparisons of HTA

Figure 2. HTA recommendations per jurisdictions as defined by the 4 assessed practices (REA, SUB, CEA, PRI).



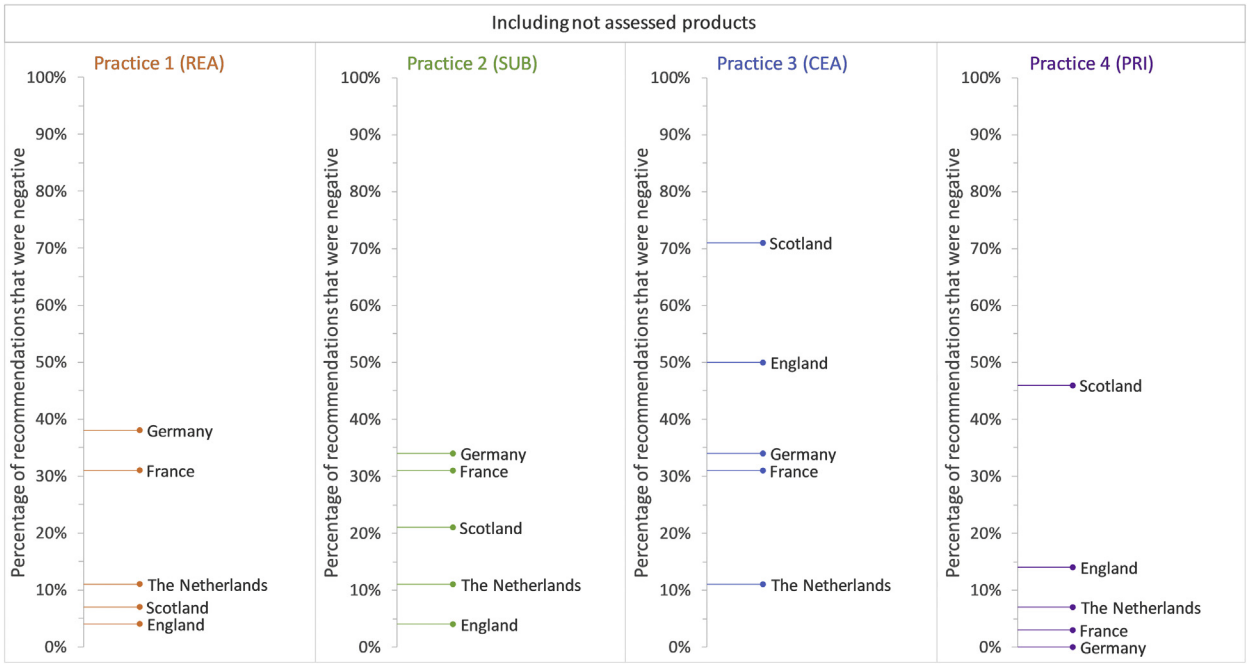
HTA indicates health technology assessment.

recommendations depends on the jurisdictions and drugs included, the scope of the research question, and other factors. Thus, a careful consideration of the possible relevance of process-related differences in HTA practices is advised.

In studies that benchmark HTA organizations or provide international multivariable models, a separate inclusion of overall outcomes (including and excluding nonsubmissions and resubmissions), REAs, and CEAs will provide part of the insight necessary to interpret the results of such studies adequately. Additionally, the aim of studies assessing HTA recommendations should be aligned with the HTA outcomes under investigation. Studies that aim to recommend on optimal HTA practices within jurisdictions are probably interested in assessing practice differences, whereas studies aiming to report on the alignment of evidence-based assessments are interested in value evaluation differences. Prediction models for HTA recommendations should adequately account for both elements. This also applies to models that try to predict HTA recommendations within a single country because different assessment practices may be available for different types of drugs in different jurisdictions. Additionally, studies trying to draw conclusions on patient access should be very careful with interpreting HTA recommendations because, in most countries, an HTA recommendation is not binding and is preceded or followed by separate processes (eg, negotiations) that ultimately affect patient access. The validity of any approach that analyses HTA recommendations depends on a valid relation between original data as stated in the HTA report and the data inputted in the analysis. Inevitably, information will be lost in this process and the institutional context of the original HTA recommendation will not be carried in full into the analysis. This is necessary to be able to compare HTA recommendations, but it also means that for valid interpretation, the results of the analysis need to be translated back to the different contexts from which the data were acquired.

HTA practices are meant to ensure an adequate balance between patient access and societal affordability (budget impact/price) within jurisdictions. To compare and evaluate HTA practices, many studies exist that investigate HTA recommendations for groups of medicines.[2-11] They differ in the extent to which they include practice-related factors as explanatory for differences between HTA jurisdictions. Based on our results, it is recommended to take a systematic approach to investigating practice differences between HTA jurisdictions. Previous research has aimed to clarify archetypes in HTA approaches based on national process maps.[22] Such archetypes may help with interpreting national HTA recommendations in an international context. Nevertheless, subsequent research has already shown that HTA jurisdictions can be classified to one archetype for one product and to another for another product, once more underlining the problems with generalizing HTA recommendations.[23] The many changes in HTA processes within jurisdictions over time and the consistent international variations in HTA practices indicate that

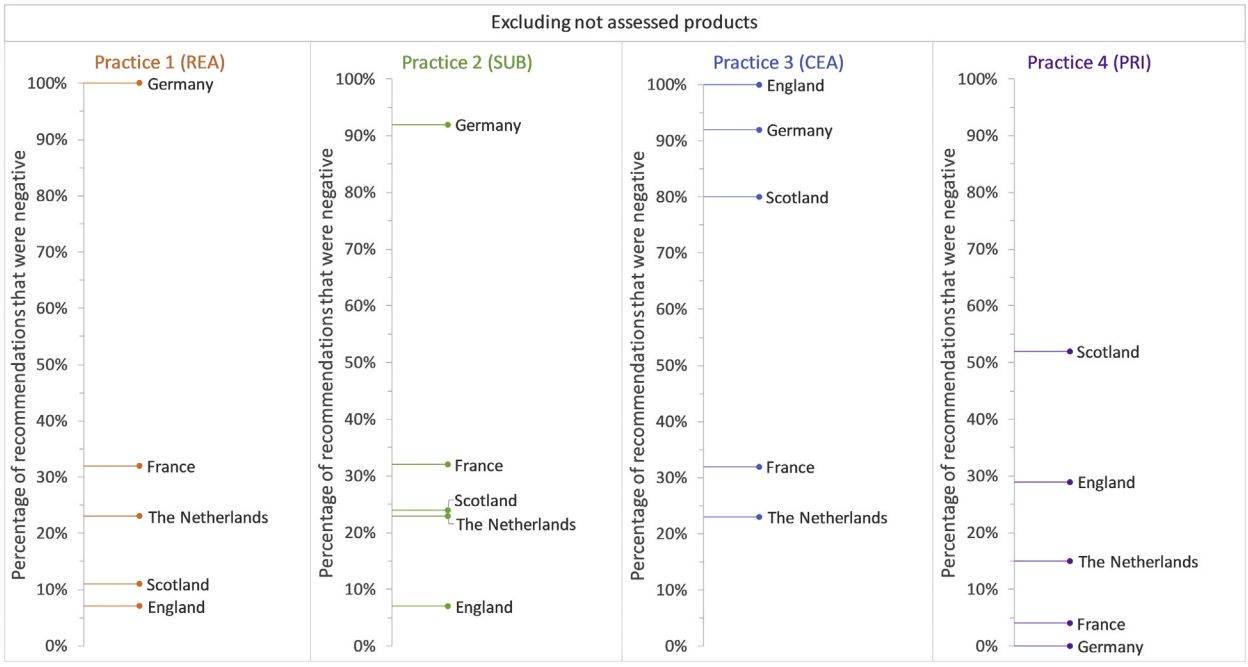
Figure 3. Proportion of negative recommendations for all jurisdictions in the 4 defined groups, including non-assessed products.



there is not yet a uniformly applicable "best practice" for HTA. Of course, different countries have different values, which can result in different preferences for reimbursement (eg, automatic reimbursement at regulatory approval or a more closed system), and HTA systems have often been shaped by political and historic arguments (eg, the UK cancer drug fund). Nevertheless, relatively similar countries might want to put extra effort into investigating best practices more thoroughly to find the most optimal processes. Differences in (timing of) patient access throughout Europe because of process-related differences between systems is an undesirable situation.

We have limited ourselves to the inclusion of CMA products because significant differences between jurisdictions had been established for these products in previous research. Although this

Figure 4. Proportion of negative recommendations for all jurisdictions in the 4 defined groups, excluding non-assessed products.



sample is relatively small, we were already able to show significant results, but the increases and reductions in negative recommendations reported in this study are not suited to be used for a general practice-correction factor. Furthermore, we have tested only four differences in HTA practices; many more may have an impact on international comparisons. For example, timing of recommendations may also be relevant because more evidence may become available over time, which can influence primary recommendations or resubmissions. Additionally, we only tested the significance of the differences between jurisdictions in proportions of negative recommendations. We did not explicitly test the significance of the changes in proportions of positive or restricted outcomes as a consequence of practice corrections; however, Figure 2 indicates that it is likely that we would find significant results on these outcomes as well. It should be noted that our data includes only five jurisdictions owing to the lack of systematically published HTA reports in other European jurisdictions and to language issues. Caution should be exercised when extrapolating our results to other jurisdictions. Finally, data extraction of individual REA or CEA results will include the interpretation of written statements in some jurisdictions, which may introduce error. One option is to have multiple researchers validate this extraction,[10] as we have done.

Conclusion

Variations in practices between international jurisdictions can have a substantial and significant impact on conclusions about recommendations by HTA bodies, as exemplified by this cohort of conditionally approved products. Studies comparing international HTA recommendations should carefully consider possible variations in practices between included jurisdictions.

Acknowledgments

The authors would like to acknowledge Lourens Bloem and Jacoline Bouvy for their role in the data extraction and validation.

Source of financial support: No funding was received for this study.

REFERENCES

1. EUnetHTA. About EunetHTA. https://www.eunethta.eu/about-eunethta/. Accessed October 31, 2018.
2. Kleijnen S, Alves TL, Meijboom K, et al. The impact of quality-of-life data in relative effectiveness assessments of new anti-cancer drugs in European countries. *Qual Life Res*. 2017;26(9):2479–2488.
3. Kleijnen S, Lipska I, Alves TL, et al. Relative effectiveness assessments of oncology medicines for pricing and reimbursement decisions in European countries. *Ann Oncol*. 2016;27(9):1768–1775.
4. Lipska I, Hoekman J, McAuslane N, Leufkens HGM, Hövels AM. Does conditional approval for new oncology drugs in Europe lead to differences in health technology assessment decisions? *Clin Pharmacol Ther*. 2015;98(5):489–491.
5. 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(2-3):167–177.
6. Nicod E, Kanavos P. Developing an evidence-based methodological framework to systematically compare HTA coverage decisions: a mixed methods study. *Health Policy*. 2016;120(1):35–45.
7. 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. *Eur J Health Econ*. 2017;18(6):715–730.
8. Spinner DS, Birt J, Walter JW, et al. Do different clinical evidence bases lead to discordant health-technology assessment decisions? An in-depth case series across three jurisdictions. *Clinicoecon Outcomes Res*. 2013;5:69–85.
9. Allen N, Walker SR, Liberti L, Salek S. Health technology assessment (HTA) case studies: factors influencing divergent HTA reimbursement recommendations in Australia, Canada, England, and Scotland. *Value Health*. 2017;20(3):320–328.
10. Vreman RA, Bouvy JC, Bloem LT, et al. Weighing of evidence by health technology assessment bodies: retrospective study of reimbursement recommendations for conditionally approved drugs. *Clin Pharmacol Ther*. 2019;105(3):684–691.
11. Akehurst RL, Abadie E, Renaudin N, Sarkozy F. Variation in health technology assessment and reimbursement processes in Europe. *Value Health*. 2017;20(1):67–76.
12. Chim L, Kelly PJ, Salkeld G, Stockler MR. Are cancer drugs less likely to be recommended for listing by the Pharmaceutical Benefits Advisory Committee in Australia? *Pharmacoeconomics*. 2010;28(6):463–475.
13. Cerri KH, Knapp M, Fernandez J-L. Decision making by NICE: examining the influences of evidence, process and context. *Health Econ Policy Law*. 2014;9(2):119–141.
14. Pauwels K, Huys I, Nys KD, Casteels M, Simoens S. Predictors for reimbursement of oncology drugs in Belgium between 2002 and 2013. *Expert Rev Pharmacoecon Outcomes Res*. 2015;15(5):859–868.
15. Dakin H, Devlin N, Feng Y, Rice N, O'Neill P, Parkin D. The influence of cost-effectiveness and other factors on Nice decisions. *Health Econ*. 2015;24(10):1256–1271.
16. Schmitz S, McCullagh L, Adams R, Barry M, Walsh C. Identifying and revealing the importance of decision-making criteria for health technology assessment: a retrospective analysis of reimbursement recommendations in Ireland. *Pharmacoeconomics*. 2016;34(9):925–937.
17. Charokopou M, Majer IM, de Raad J, Broekhuizen S, Postma M, Heeg B. Which factors enhance positive drug reimbursement recommendation in Scotland? A retrospective analysis 2006-2013. *Value Health*. 2015;18(2):284–291.
18. Cerri KH, Knapp M, Fernandez J-L. Public funding of pharmaceuticals in the Netherlands: investigating the effect of evidence, process and context on CVZ decision-making. *Eur J Health Econ*. 2014;15(7):681–695.
19. Salas-Vega S, Bertling A, Mossialos E. A comparative study of drug listing recommendations and the decision-making process in Australia, the Netherlands, Sweden, and the UK. *Health Policy*. 2016;120(10):1104–1114.
20. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)*. 2012;22(3):276–282.
21. Warrens MJ. Equivalences of weighted kappas for multiple raters. *Statistical Methodology*. 2012;9(3):407–422.
22. Allen N, Pichler F, Wang T, Patel S, Salek S. Development of archetypes for non-ranking classification and comparison of European National Health Technology Assessment systems. *Health Policy*. 2013;113(3):305–312.
23. Allen N, Liberti L, Walker SR, Salek S. A comparison of reimbursement recommendations by European HTA agencies: is there opportunity for further alignment? *Front Pharmacol*. 2017;8:384.