ELSEVIER

feature

European regulatory use and impact of subgroup evaluation in marketing authorisation applications

Julien Tanniou[1,2], j.tanniou@umcutrecht.nl, Steven Teerenstra[2,3], Sagal Hassan[2], Andre Elferink[2], Ingeborg van der Tweel[1], Christine Gispen-de Wied[2] and Kit C.B. Roes[1,2]

Marketing authorisation application dossiers relating to medicinal products containing new active substances and evaluated by the European Medicines Agency (EMA) over the period 2012–2015 were examined. Major objections and other concerns relating to efficacy and safety of the day 80 assessment reports were reviewed. Overall, approved products have more subgroup concerns than nonapproved products, which seems to be a consistent pattern. Subgroup analyses are mainly assessed to have the insurance that subgroups of patients that might lack a positive benefit: risk ratio will not be wrongly included in the approved treatment indication.

Introduction

Before a medicine can be sold or prescribed to citizens across the European Union (EU), a marketing authorisation must be obtained. The European regulation offers several options for the authorisation of medicinal products: the centralised procedure, the mutual recognition procedure, the decentralised procedure and the national procedure. Today, the majority of new, innovative medicines passes through the centralised procedure to be marketed in the EU, with the objective to ensure their efficacy, safety and quality. Confirmatory (pivotal) clinical trials are usually performed to inform a benefit–risk decision, the results of which will be the basis for a treatment recommendation (labelling). It is well recognised that the balance of benefits and risks can vary across the patient population [1–4]. Therefore, subgroup analyses constitute a fundamental step in the assessment of a marketing authorisation application (MAA) so as to make optimal decisions at the population level and for each patient. As recently outlined by the draft guideline on the investigation of subgroups in confirmatory clinical trials [5], the role of subgroup analyses might differ depending on the overall results of the trial(s). If the presented clinical data are convincing overall, this general trend should be confirmed across subgroups of clinical importance. By contrast, if the overall results are borderline (positive) it might be of interest to identify a subgroup with persuasive results. These investigations have potentially important consequences for the medicinal product licensing, labelling, reimbursement and treatment decisions, when results from analyses of the overall (pooled) trial population might not hold for important subpopulations. It is therefore of interest to investigate how the regulatory evaluation of subgroups impacts MAAs.

Regulatory evaluation of subgroups over the period 2012–2015

All MAAs of new active substances (NAS) evaluated by the EMA through the centralised procedure between 1 January 2012 and 31 December 2015 were included in the study. The Dutch Medicines Evaluation Board (MEB) annual reports (Appendix A) have been used to retrieve the NAS status of each approved application [6–9]. With respect to the nonapproved products, the EMA website was consulted to obtain the list of 'refused' products [10]. Because the number of nonapproved products was rather small, all applications withdrawn before final Committee for Medicinal Products for Human Use (CHMP) decision on marketing authorisation were also

http://dx.doi.org/10.1016/j.drudis.2017.09.012

taken into account [11]. Each European Public Assessment Report (EPAR) was examined to see whether the active substance was originally considered as a NAS.

In the European assessment procedure, two member-states are appointed to take the lead: their respective CHMP members are then the so-called rapporteur and co-rapporteur. The first preliminary assessment reports from rapporteur and co-rapporteur are sent to the applicant at day 80 (i.e., 80 days after the official start of the procedure). In the day 80 reports, a specific section is dedicated to a 'list of questions' with a subsection about 'clinical aspects' which contains distinct efficacy and safety parts. Two types of criticisms: major objections (MOs) and other concerns (OCs), are reported. The MOs and OCs relating to efficacy and safety of the day 80 reports constitute the raw data of this study, because these reports were considered the most exhaustive available regarding potential issues raised. We identified all MOs and OCs related to subgroup evaluations. We defined a broad automatic (text mining) search strategy, performed by SH, based on various keywords: sub(-)group(s), sub(-)population(s), sub(-)set(s), mutation(s) or marker(s). Because MOs are the most relevant criticisms, J.T. read all efficacy and safety MOs to be reassured not to miss any important objections related to subgroup assessments. All selected MOs and OCs were reviewed by J.T. and S.T. to decide which should be retained for the analysis. As illustrated by their nonconsideration in the draft guideline on the investigation of subgroups in confirmatory clinical trials [5], subgroup MOs/OCs related to either treatment regimen (doses, duration, etc.), subgroups based on post-randomisation variables (e.g., the subgroup of responders) and subgroups for which the concern was a lack of data (e.g., elderly) or were the possibility/validity of extrapolation to that subgroup not retained as not directly related to subgroup analysis from a statistical or methodological point of view. All MOs and OCs were classified to one of the three categories: (i) consistency or heterogeneity of the subgroup results compared to the overall result; (ii) proposal to search for a subgroup with better efficacy and/or better safety; or (iii) statement that the indication should be restricted to a subgroup. J.T. and S.T. first performed this task independently. When J.T. and S.T. had a divergent opinion they discussed to reach a final agreement.

For each approved application, the proposed indication by the applicant as well as the final approved indication were compared to assess whether a change in the indication had been observed. J.T. and A.E. independently made their own classification to decide on whether the proposed indication and the approved indication were similar. They also compared their results and, when a divergent opinion arose, a final agreement was made between these authors. All results presented are provided in contingency tables containing absolute and relative numbers. These results are counts at the application level (i.e., the number of applications with, for example, at least one efficacy MO related to subgroup evaluation). The same holds for efficacy OCs, safety MOs and safety OCs. Finally, a distinction has been made between approved (with or without restriction of indication) and nonapproved applications, orphan and non-orphan status.

Review of subgroup assessments in MAAs

According to the Dutch MEB annual reports, there were 43 authorised medicinal products for human use with a NAS status in 2015, 35 in 2014, 47 in 2013 and 26 in 2012. A limited number of products had to be removed (13) because they either could not be found in the database or were considered as known active substance in their respective EPARs. The dataset is therefore composed of 138 authorised applications. In total, 15 applications were refused. By looking at each EPAR individually, we found that one product was a generic and four were considered to be known active substances. These five were therefore not retained in our dataset. Nine requested the active substance to be considered as a NAS, but the EMA CHMP was of the opinion that it was not appropriate to conclude on the NAS status at that time, in light of the negative recommendation, and one was qualified as a NAS by the CHMP. Following the applicant's original request or CHMP qualification regarding NAS, we decided to keep these ten refused applications. Because the number of refused applications was limited, we considered the applications that were withdrawn by the applicant before final CHMP decision on marketing authorisation. Between 2012 and 2015 there were 39 such applications, of which 25 were considered as known active substances and were therefore not retained. Eleven applications were qualified as a NAS and three requested the active substance to be considered as a NAS but did not receive an answer at the time of the withdrawal. These 14 applications were kept for analysis. All retained applications (authorised, refused and withdrawn before final CHMP decision on marketing authorisation) are listed (see Appendix A in Supplementary material, available online).

According to Table 1, subgroup MOs/OCs are often present in day 80 reports (68%). They are however more prominent in authorised applications than in refused/withdrawn applications (70% vs 58%). Regarding efficacy MOs, efficacy OCs and safety OCs, subgroup criticisms are approximately twice as frequent in authorised applications as in refused/withdrawn applications. By contrast, safety MOs are more common in refused/withdrawn applications than in authorised applications (21% vs 9%).

Regarding orphan medicinal products, almost no difference is observed (24% vs 29%) regarding efficacy MOs between authorised and refused/withdrawn applications, whereas authorised applications have more safety MOs than refused/withdrawn applications (9% vs 0%) (Table 2). Concerning non-orphan medicinal products, the difference between efficacy MOs is more pronounced (25% vs 6%) in favour of authorised applications, even though there is an opposite effect with more safety MOs in refused/withdrawn applications (10% vs 29%).

TABLE 1

Subgroup MOs/OCs in authorised and refused/withdrawn applications

	Authorised applications (*n* = 138)	Refused/withdrawn applications (*n* = 24)	All applications(*n* = 162)
At least one efficacy MO	34 (25%)	3 (12%)	37 (23%)
At least one efficacy OC	82 (59%)	8 (33%)	90 (56%)
At least one safety MO	13 (9%)	5 (21%)	18 (11%)
At least one safety OC	31 (22%)	3 (12%)	34 (21%)
Any MO or OC	96 (70%)	14 (58%)	110 (68%)

TABLE 2

Subgroups MOs/OCs in authorised and refused/withdrawn applications depending on the orphan status

	Authorised applications (*n* = 138)		Refused/withdrawn applications (*n* = 24)	
	Orphan status (n = 33)	*Non-orphan status (n = 105)*	*Orphan status (n = 7)*	*Non-orphan status (n = 17)*
At least one efficacy MO	8 (24%)	26 (25%)	2 (29%)	1 (6%)
At least one efficacy OC	24 (73%)	58 (55%)	1 (14%)	7 (41%)
At least one safety MO	3 (9%)	10 (10%)	0 (0%)	5 (29%)
At least one safety OC	7 (21%)	24 (23%)	0 (0%)	3 (18%)
Any MO or OC	26 (79%)	70 (67%)	3 (43%)	11 (65%)

Among all MOs and OCs related to subgroup evaluations, we observe the general pattern that they are mostly dedicated to the assessment of consistency or heterogeneity of the overall treatment effect. The related extensive table is provided (see supplementary material online, Table S1). For instance, out of 34 authorised applications with at least one subgroup related efficacy MO, 27 were classified as consistency or heterogeneity of the overall result across subgroups, six as proposal to search for a subgroup with better benefit–risk characteristics and two as statement that the indication should be restricted to a subgroup. Please note that, because one application can have several efficacy MOs, this application might be classified in more than one category – the reason why it does not add up to 34 in this case. To clarify the underlying issues raised, we provide illustrative examples for each category. To ensure confidentiality, some words and/or sentences have been removed and/or replaced without substantially changing the content.

Consistency or heterogeneity of the subgroup results compared to the overall result

- The results were inconsistent across important subgroups such as gender (marginal effect in female patients, who are generally more affected) [. . .]. The Applicant should comment on these inconsistencies.
- Subgroup analyses of progression-free survival (PFS) in subjects with an early stage disease at baseline did not show any benefit [. . .]. The clinical relevance of the treatment in this subpopulation is therefore debatable. Please discuss.

Proposal to search for a subgroup with better efficacy and/or better safety

- Although activity of treatment appears to be demonstrated, the magnitude of the effect observed appears to be clearly inferior to other standard treatment options currently available [. . .]. The company should discuss and justify for which patients treatment could have a positive benefit:risk.

Statement that the indication should be restricted to a subgroup

- Treatment cannot be recommended to patients with moderate disease. Efficacy of treatment in these patients is decreased with a higher frequency of adverse events. Therefore, in patients with moderate disease, the benefit:risk ratio is considered negative.

TABLE 3

Subgroup MOs/OCs in authorised applications with and without change of indication

	Different indication (*n* = 50)	Same indication (*n* = 88)
At least one efficacy MO	17 (34%)	17 (19%)
At least one efficacy OC	30 (60%)	52 (59%)
At least one safety MO	3 (6%)	10 (11%)
At least one safety OC	17 (34%)	14 (16%)
Any MO or OC	37 (74%)	59 (67%)

Because we noted that subgroup MOs/OCs are more frequent in authorised applications than in refused/withdrawn applications (Table 1), we therefore looked more in-depth into these authorised applications. We compared the subgroup MOs/OCs in authorised applications to see whether the original proposed indication was similar to or different from the final approved indication (Table 3).

Among the 138 authorised applications, 88 were approved without any substantial changes of indication, and 50 obtained an approval with a restriction of indication. Authorised applications with a change of indication tend to have slightly more subgroup MOs/OCs (74% vs 67%). Looking at Table 3, there is no (large) difference for efficacy OCs

TABLE 4

Subgroup MOs/OCs in authorised applications with and without change of indication depending on the orphan status

	Different indication (*n* = 50)		Same indication (*n* = 88)	
	Orphan status (n = 16)	*Non-orphan status (n = 34)*	*Orphan status (n = 17)*	*Non-orphan status (n = 71)*
At least one efficacy MO	4 (25%)	13 (38%)	4 (24%)	13 (18%)
At least one efficacy OC	11 (69%)	19 (56%)	13 (76%)	39 (55%)
At least one safety MO	1 (6%)	2 (6%)	2 (12%)	8 (11%)
At least one safety OC	3 (19%)	14 (41%)	4 (24%)	10 (14%)
Any MO or OC	12 (75%)	25 (74%)	14 (82%)	45 (63%)

(60% vs 59%) and safety MOs (6% vs 11%). Conversely, authorised applications with different initial and final indications more often have efficacy MOs (34% vs 19%) and more safety OCs (34% vs 16%).

Depending on the orphan status, we clearly see a difference concerning efficacy MOs (Table 4). Regarding orphan medicinal products, almost no difference is observed (25% vs 24%), whereas for non-orphan medicial products the difference between authorised applications with and without a change of indication is substantial (38% vs 18%). The same holds for safety OCs (orphan medicinal products: 19% vs 24%; non-orphan medicinal products: 41% vs 14%). As mentioned previously, most of the MOs/OCs are again related to consistency or heterogeneity (see supplementary material online, Table S2).

Concluding remarks

In line with the recent EMA draft guideline on the investigation of subgroups in confirmatory clinical trials [5], it appears that subgroup analyses are an integral part of clinical trial planning, analysis and inference during the assessment of a MAA. They can either reinforce or contradict the overall results, hence having a direct influence on the medicinal product licensing, labelling, reimbursement and prescribing decisions. In this study, we note that subgroup-related MOs and/or OCs are prominent (68%) in day 80 reports over the period 2012–2015. This result emphasises the essential role of subgroups in the assessment of MAAs. Consequently, the investigation on how subgroups were regulatory assessed in recent MAAs is of significance.

Interestingly, we observed that efficacy MOs are more frequent in authorised applications, whereas safety MOs are more common in refused/withdrawn applications. We could therefore imagine that in refused/withdrawn applications the safety concerns dominate the decision. By contrast, in authorised applications, investigating the consistency of treatment effect across subgroups of clinical importance is a fundamental step in the assessment process. Refused/withdrawn applications probably face more-important issues than the consistency of treatment effect across subgroups, which are of less interest as long as the main issues are not solved. We noticed that the difference between authorised and refused/withdrawn applications does not exist with orphan medicinal products, but is clearly present with non-orphan drugs. This could suggest that more subgroup analyses are requested when data are not sparse. We also investigated whether subgroups impact the choice of a medicine's final indication. We noticed that MAAs with a change of indication had substantially more efficacy MOs than MAAs without a change of indication in situations when the investigated drug is not designated as an orphan medicinal product. This distinction does not exist for orphan drugs. Moreover, these efficacy MOs are mainly related to the consistency or heterogeneity of the treatment effect. Subgroup analyses are therefore mainly assessed to have the insurance that subgroups will not be wrongly included in the final indication.

Although not the purpose of this study, it is interesting to note that, among the 162 reviewed applications, the UK followed by Sweden, Germany and The Netherlands were responsible for 186 (57%) out of the 324 (co-) rapporteurships. The fact that these countries do most of the centralised procedures of NAS applications reflects not so much the size of the respective countries but rather the strategy and priorities of the national regulatory agencies. Even though the majority of assessment work is done by a minority of countries, there is no indication that the rapporteur country influenced the subgroup MO/OC or indication changes (logistic regression analysis not shown), which provides a reassuringly consistent assessment about subgroups across national competent authorities.

Despite the pivotal role of subgroup analyses in regulatory assessments, this study is, to our knowledge, the first to investigate how European regulators deal with subgroups during the procedure of obtaining a marketing authorisation. The main limitation of this study is the small number of refused applications. Combining 'withdrawn before final CHMP decision' on marketing authorisation and refused applications might not be an ideal solution but is the best surrogate we could think of to minimise this limitation. It should therefore be realised that estimated percentages are less precise, mainly for the refused/withdrawn applications. The focus of this study was on methodological aspects of subgroup analyses in MAAs. However, multiple other (nonstatistical) considerations would influence whether or when issues related to subgroups are raised and their impact in terms of regulatory decisions (e.g., approvability or labelling of a medicinal product). Subgroup analyses are known to be prone to statistical and methodological issues such as inflation of type I error owing to multiple testing, low power, inappropriate statistical analyses or lack of prespecification [12–17]. Given that a close inspection of relevant subgroups is important and performed in practice, regulators should clearly be aware of these issues as well as their potential respective solutions [18].

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.drudis.2017.09.012.

References

1 Grouin, J.M. *et al.* (2005) Subgroup analyses in randomized clinical trials: statistical and regulatory issues. *J. Biopharm. Stat.* 15, 869–882

2 Maggioni, A.P. *et al.* (2007) FDA and CPMP rulings on subgroup analyses. *Cardiology* 107, 97–102

3 Hemmings, R. (2014) An overview of statistical and regulatory issues in the planning, analysis, and interpretation of subgroup analyses in confirmatory clinical trials. *J. Biopharm. Stat.* 24, 4–18

4 Koch, A. and Framke, T. (2014) Reliably basing conclusions on subgroups of randomized clinical trials. *J. Biopharm. Stat.* 24, 42–57

5 EMA/CHMP/539146/2013 Committee for Medicinal Products for Human Use CHMP (2014) *Guideline on the Investigation of Subgroups in Confirmatory Clinical Trials.* Available at: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2014/02/WC500160523.pdf

6 MEB Annual Report (2012) Available at: https://english.cbg-meb.nl/documents/annual-reports/2012/01/01/annual-report-2012.

7 MEB Annual Report (2013) Available at: https://english.cbg-meb.nl/about-meb/documents/annual-reports/2013/01/01/annual-report-201.

8 MEB Annual Report (2014) Available at: https://english.cbg-meb.nl/about-meb/documents/annual-reports/2014/01/01/annual-report-2014.

9 MEB Annual Report (2015) Available at: https://english.cbg-meb.nl/about-meb/documents/annual-reports/2015/01/01/annual-report-2015.

10 List of Refused Products. Available at: http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/landing/epar_search.jsp&mid=WC0b01ac058001d124&startLetter=View%20all&searchTab=&keyword=Enter%20keywords&searchType=name&alreadyLoaded=true&status=Refused&jsenabled=false&searchGenericType=generics&orderBy=authDate&pageNo=1

11 List of Withdrawn Before Approval Products. Available at: http://www.ema.europa.eu/ema/index.jsp?searchType=Name&type=Initial+authorisation&mid=WC0b01ac058001d128&curl=pages%2Fmedicines%2Flanding%2Fwapp_search.jsp&keyword=Enter+keywords&alreadyLoaded=true&startLetter=View+all&pageNo=1

12 Oxman, A.D. and Guyatt, G.H. (1992) A consumer's guide to subgroup analyses. *Ann. Intern. Med.* 116, 78–84

13 Rothwell, P.M. (2005) Treating individuals 2. Subgroup analysis in randomized controlled trials: importance, indications, and interpretation. *Lancet* 365, 176–186

14 Dijkman, B. *et al.* (2009) How to work with a subgroup analysis. *Can. J. Surg.* 52, 515–522

15 Sun, X. *et al.* (2010) Is a subgroup effect believable? Updating criteria to evaluate the credibility of subgroup analyses. *BMJ* 340, c117
16 Sun, X. *et al.* (2011) Is a subgroup claim believable? A user's guide to subgroup analyses in the surgical literature. *J. Bone Joint Surg. Am.* 93, e8
17 Sun, X. *et al.* (2012) Credibility of claims of subgroup effects in randomised controlled trials: systematic review. *BMJ* 344, e1553
18 Tanniou, J. *et al.* (2016) Subgroup analyses in confirmatory clinical trials: time to be specific about their purposes. *BMC Med. Res. Methodol.* 16 (Feburay (18)), 20http://dx.doi.org/10.1186/s12874-016-0122-6

Julien Tanniou[1,2,*]
Steven Teerenstra[2,3]
Sagal Hassan[2]
Andre Elferink[2]
Ingeborg van der Tweel[1]
Christine Gispen-de Wied[2]
Kit C.B. Roes[1,2]

[1]*Julius Center for Health Sciences and Primary Care, University Medical Centre, Utrecht, The Netherlands*

[2]*Medicines Evaluation Board, Utrecht, The Netherlands*
[3]*Radboud Institute for Health Sciences, Department of Health Evidence, Section Biostatistics, Radboud University Medical Centre, Nijmegen, The Netherlands*

**Corresponding author.*