Original research

Attitudes of healthcare professionals and drug regulators about progression-free survival as endpoint in the advanced cancer setting

Douwe Postmus [a,b], Saskia Litiere [c], Jan Bogaerts [c], Jurjen Versluis [d], Jan J. Cornelissen [a,d], Francesco Pignatti [a,*]

[a] European Medicines Agency, Amsterdam, the Netherlands
[b] University of Groningen, University Medical Center Groningen, Department of Epidemiology, Groningen, the Netherlands
[c] European Organisation for Research and Treatment of Cancer, Brussels, Belgium
[d] Erasmus Medical Center Cancer Institute, Department of Hematology, Rotterdam, the Netherlands

ARTICLE INFO

ABSTRACT

Purpose: To describe the attitudes of healthcare professionals and drug regulators about progression-free survival (PFS) as efficacy endpoint in clinical trials with patients with advanced cancer and to explore to what extent these attitudes influence the willingness to trade between PFS and toxicity.
Methods: Cross-sectional survey with regulators from the European Medicines Agency (EMA), and healthcare professionals (HCP) from the "Stichting Hemato-Oncologie voor Volwassenen Nederland" (HOVON) collaborative group and the European Organisation for Research and Treatment of Cancer (EORTC). Attitudes towards PFS were elicited using 5-point Likert items. The respondents' willingness to trade between PFS and grade 3 or 4 (G34) toxicity was assessed using the threshold technique and quantified in terms of their maximum acceptable risk (MAR).
Results: Responses were collected from 287 HCPs and 64 regulators with mainly clinical expertise. Attitudes towards PFS were often spread out in both groups and related to beliefs about PFS being a likely surrogate for clinical benefit, being an intrinsic benefit to be distinguished from OS, or on the importance given to OS. Being a regulator or holding stronger beliefs about PFS being a likely surrogate or an intrinsic benefit were associated with a higher MAR. Presence of a supportive trend in OS was stated as important but was not associated with MAR. There was agreement on the need to address bias in the adjudication of PFS and the need for improving communication to patients about meaning, strengths, and limitations of improvements in PFS.
Conclusion: Attitudes towards PFS were spread out and were associated with individual differences in the willingness to trade between toxicity and PFS. There was agreement on the need to address bias in the adjudication of PFS and improving communication to patients.

1. Introduction

The clinical importance of different endpoints in assessing cancer drugs has often been a matter of debate among drug developers, regulators, clinicians, and patients. While most agree that effects in terms of overall survival (OS) dominate all others, there are varied opinions as to the value of progression-free survival (PFS) [1,2]. Some, highlight its timeliness; the fact that unlike OS it is unaffected by subsequent treatments; and that it is likely related to the onset or worsening of symptoms and the need for subsequent less effective or more toxic treatments [3]. Others, stress the radiographic rather than clinical nature of the

endpoint; the associated costs, error, and potential bias in adjudication; the lack of surrogacy for important effects like OS and health-related quality of life (HRQoL); and that it fosters false hopes and diverts important resources to the search for "true" benefits [4].

When PFS is the main efficacy endpoint in the pivotal clinical trial of a new cancer drug, differences in individual attitudes about the use of this endpoint may translate into different opinions about the benefit-harm balance. The impact of different attitudes may be described in terms of the willingness to trade between toxicity and PFS. Substitution rates between PFS and toxicity may also vary depending on the size of the PFS effect and amount of supportive evidence, especially from

---

analysis of OS. For example, willingness to trade toxicity for PFS may be higher if the effects on PFS are larger and if there is a positive trend on OS. Currently, the impact of attitudes about PFS and presence of supportive OS trends on trade-offs between toxicity and PFS is not known.

The aims of this study were to describe the distribution of attitudes of healthcare professionals (HCPs) and drug regulators towards the use of PFS as a primary efficacy endpoint for the assessment of treatments for advanced cancer, and to describe trade-offs between toxicity and PFS, with or without supportive OS trends.

2. Methods

2.1. Study design and participants

An online, cross-sectional survey was designed to be conducted among drug regulators and HCPs. The regulators' group included assessors who had either participated in the scientific evaluation of the clinical part of oncology drug applications evaluated by EMA in 2019–2021 or were part of EMA's Oncology European Specialised Expert Community, including assessors and experts that contribute to EMA's scientific work as members of scientific committees, working parties, or drafting groups. The HCPgroup consisted of investigators and researchers in either the "Stichting Hemato-Oncologie voor Volwassenen Nederland" (HOVON) collaborative group or the European Organisation for Research and Treatment of Cancer (EORTC) network. Email invitations with a public link to the survey were sent to prospective study participants. The survey was open from 18 May 2022 to 5 July 2022 and responses were collected anonymously.

2.2. Questionnaire

The questionnaire was in English, and it was created in the web application Research Electronic Data Capture 10.0.23 (REDCap. URL www.projectredcap.org). It consisted of four parts: questions about professional role and time in role; questions about potential biases associated with the use of PFS as a main efficacy endpoint; questions about the relationship between PFS and other efficacy endpoints in a general advanced cancer setting; and questions about the willingness to trade between PFS and severe or life-threatening (G34) toxicity risk. All parts were presented in the context of well-conducted randomised controlled trials with PFS as primary endpoint. The questions in the parts about potential biases and the relationship between PFS and other efficacy endpoints consisted of 5-point Likert scale items constructed based on expert input from the study team. The participants' willingness to trade between PFS and the risk of G34 toxicity was assessed using the threshold technique [5]. Participants were randomised to one of two threshold exercises with identical treatment effects on PFS and G34 toxicity but with different results of the supportive OS analysis: a scenario where median OS was not reached in both study arms accompanied by the statement that there was no apparent detrimental effect based on visual exploration (subsequently referred to as the "no detrimental effect" scenario), and a scenario where there was a difference in median OS of 3 months accompanied by the statement that the effect is almost but not quite statistically significant ("positive trend scenario").

The question tree used for the two threshold exercises is schematically depicted in Supplementary Fig 1. In short, both exercises started with an initial choice between two hypothetical treatments that was the same to all participants. Depending on a participant's response to this first question, either the median PFS or the proportion of patients experiencing a G34 event associated with the experimental treatment was decreased. In the subsequent questions, further changes to the treatment outcome in that cell of the table were made until a participant was indifferent between the two treatment options or a maximum number of choice questions was reached (up to 4 questions depending on the path taken).

2.3. Statistical analysis

The responses to the potential biases associated with the use of PFS questions and the relationship between PFS and other efficacy endpoints questions were summarised graphically. The Mann-Whitney U test was used to test for differences between groups.

The correlation between individual responses to questions about the relationship between PFS and other efficacy endpoints were explored with factor analysis. The number of latent factors, each unveiling a different aspect of this relationship, was determined using the Bayesian Information Criterion. Participants' latent factor scores were estimated using Thurstone's least squares regression. Associations between estimated factor scores and participants' current role were examined using linear regression analysis.

Based on the possible paths through the question tree for the thresholding exercise, the participants' willingness to trade between PFS and the risk of G34 toxicity was classified into seven maximum acceptable risk (MAR) categories. Mann-Whitney U tests were used to test for differences in the distribution of these MAR categories between HCPs and regulators, as well as between the two OS outcome scenarios. Additionally, Kruskal-Wallis tests were employed to examine differences in the MAR distributions among subgroups defined by tertiles of estimated factor scores on each factor.

P-values <0.05 were considered as statistically significant. All analyses were conducted in R version 4.2.3 (R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/).

3. Results

3.1. Participants' characteristics

The survey was sent to 443, 2639, and 336 separate email addresses for the regulators, EORTC, and HOVON, respectively. The survey was opened 395 times. After deletion of the blank records, 351 responses (92 partially completed and 259 fully completed) remained. Out of these remaining responses, 287 (81.8%) were provided by HCPs and 64 (18.2%) by regulators. Within the HCP group, 130 (45.3%) were oncologists, 95 (33.1%) were haematologists, 53 (18.5%) were physicians from other specialties, and 9 (3.1%) were other HCPs, health scientists, or patient advocates. Within the regulators' group, 56 (87.5%) were clinical regulators. Time in role was longer in the HCP group than in the regulators' group, with 70.0% of the HCPs having spent longer than 10 years in their current role vs 34.4% in the regulator group. Both groups expressed a good familiarity with the design, conduct, and analysis of randomised controlled trials in the clinical development of treatment of advanced cancer as well as with balancing the benefits and harms of cancer treatments (Table 1).

3.2. Biases associated with the use of PFS

Participants' views about potential biases associated with PFS as efficacy endpoint are summarised graphically in Fig. 1 and numerically in Supplementary Table 1. A large proportion of respondents (31.4% to 49.0% within each group) believed that treatment effect estimates are very or extremely likely biased in open-label trials due to unbalanced assessment or reader evaluation bias, or that insufficient attention is being paid to informative censoring. The majority agreed on the importance of sensitivity analyses, of strict control within the study protocol to prevent informative censoring, and especially on the importance of detailed description of reasons for all cases of early censoring.

3.3. Views on PFS in relation to other endpoints

Respondents' views on the relation between PFS and other endpoints in a general advanced cancer setting are summarised graphically in

Table 1
Characteristics of the study population.

	Healthcare professionals (n = 287)	Regulators (n = 64)
Current role		
Doctor - oncologist	130 (45.3%)	
Doctor - haematologist	95 (33.1%)	
Doctor – other	53 (18.5%)	
Other healthcare professional	4 (1.4%)	
Other scientist	3 (1%)	
Patient advocate	2 (0.7%)	
Regulator - clinical		56 (87.5%)
Regulator - other		8 (12.5%)
Time in role		
0 - 5 years	30 (10.5%)	18 (28.1%)
5 - 10 years	56 (19.5%)	24 (37.5%)
>10 years	201 (70%)	22 (34.4%)
Familiarity with the design, analysis, and conduct of randomised clinical trials in the clinical development of treatments of advanced cancer		
Extremely	55 (19.2%)	13 (20.3%)
Very	148 (51.6%)	28 (43.8%)
Moderately	68 (23.7%)	20 (31.2%)
Slightly	10 (3.5%)	2 (3.1%)
Not at all	6 (2.1%)	1 (1.6%)
Familiarity with balancing the benefits and harms of cancer treatments		
Extremely	84 (29.3%)	15 (23.4%)
Very	160 (55.7%)	29 (45.3%)
Moderately	33 (11.5%)	17 (26.6%)
Slightly	6 (2.1%)	2 (3.1%)
Not at all	4 (1.4%)	1 (1.6%)

Fig. 2 and numerically in Supplementary Table 2. The majority view was that the surrogacy of PFS for either OS or HRQoL is infrequent and that improvements in PFS should come with at least a favourable trend in OS. A small subgroup (16.5% and 6.0%, for HCP and regulators, respectively), agreed or strongly agreed that improvements in PFS could still be convincing even if there was a negative trend in OS. Insistence on proven OS benefits was believed to delay access to innovative treatments by the majority of HCPs (59.7% and 37.3% for HCPs and regulators, respectively) but also to encourage the development of truly promising compounds (61.5% and 75.0%, respectively). Respondents most frequently agreed that PFS is the preferred endpoint in first-line trials (65.5% and 68.0% for HCPs and regulators, respectively).

The majority in both groups disagreed, often strongly for HCPs, that meaning, strengths, and limitations of improvements in PFS are well-understood by patients (76.9% and 61.3% for HCPs and regulators, respectively).

According to the results of the exploratory factor analysis, the correlation among the individual responses to these eight Likert items could be explained in terms of three underlying factors (Fig. 3). The three factors could be interpreted as: capturing the respondents' views towards PFS as a likely surrogate for OS and HRQoL (factor 1); PFS as a benefit different from OS (factor 2); and the importance of OS (factor 3). No statistical differences were found in the mean factor scores between HCPs and regulators on the first two factors. For the third factor, the mean score was 0.41 standard deviations (SDs) higher for the regulators group compared to the HCP group (p-value = 0.013).

Within the HCP group, the mean score on the first factor was 0.40 SDs higher for oncologists compared to haematologists (p = 0.02). No statistical differences were found between oncologists, haematologists, and other HCPs with respect to the mean scores on the other two factors. Within the regulators' group, there were no statistical differences between clinical and other regulators with respect to the mean scores on any of the three factors.
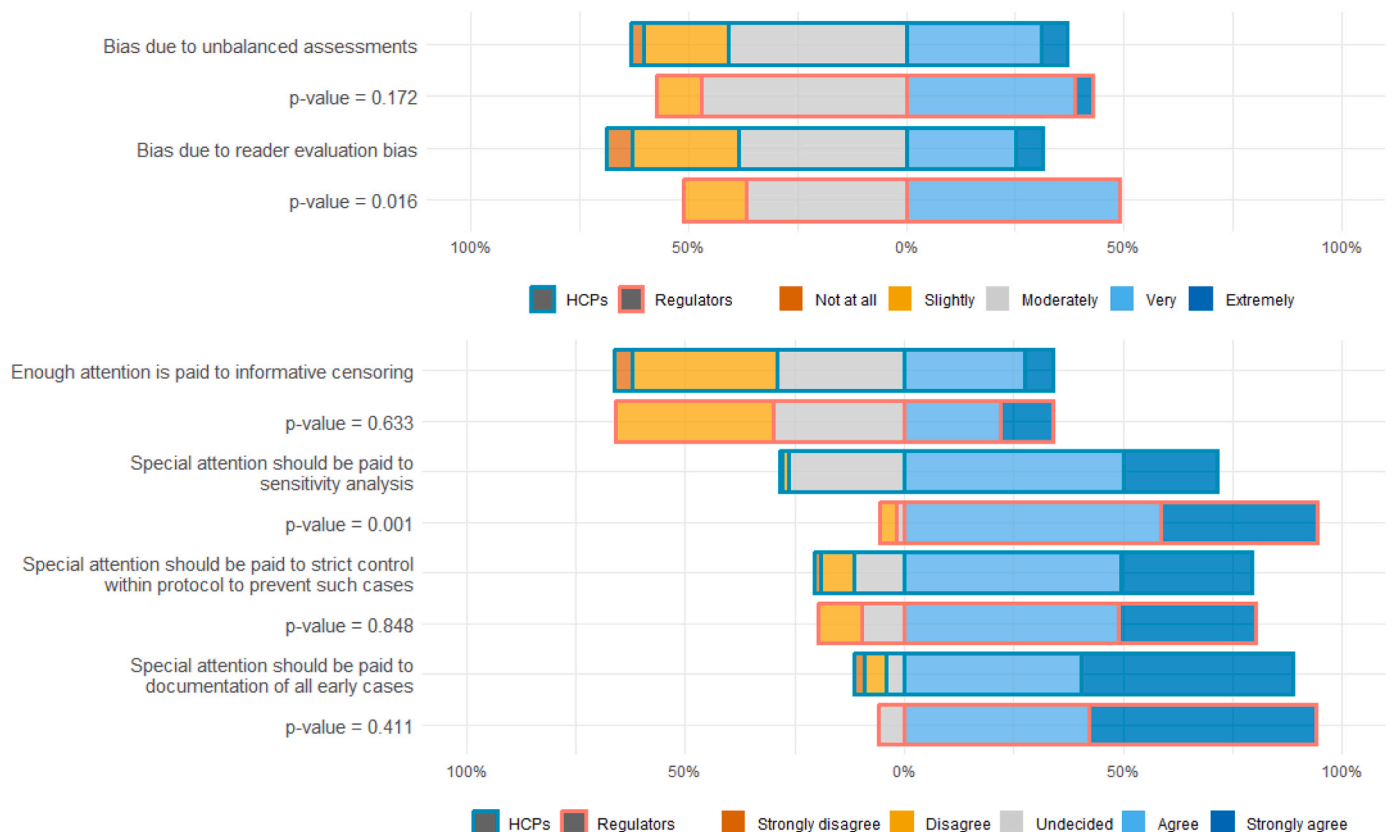
Fig. 1. Responses to potential biases associated with PFS questions.
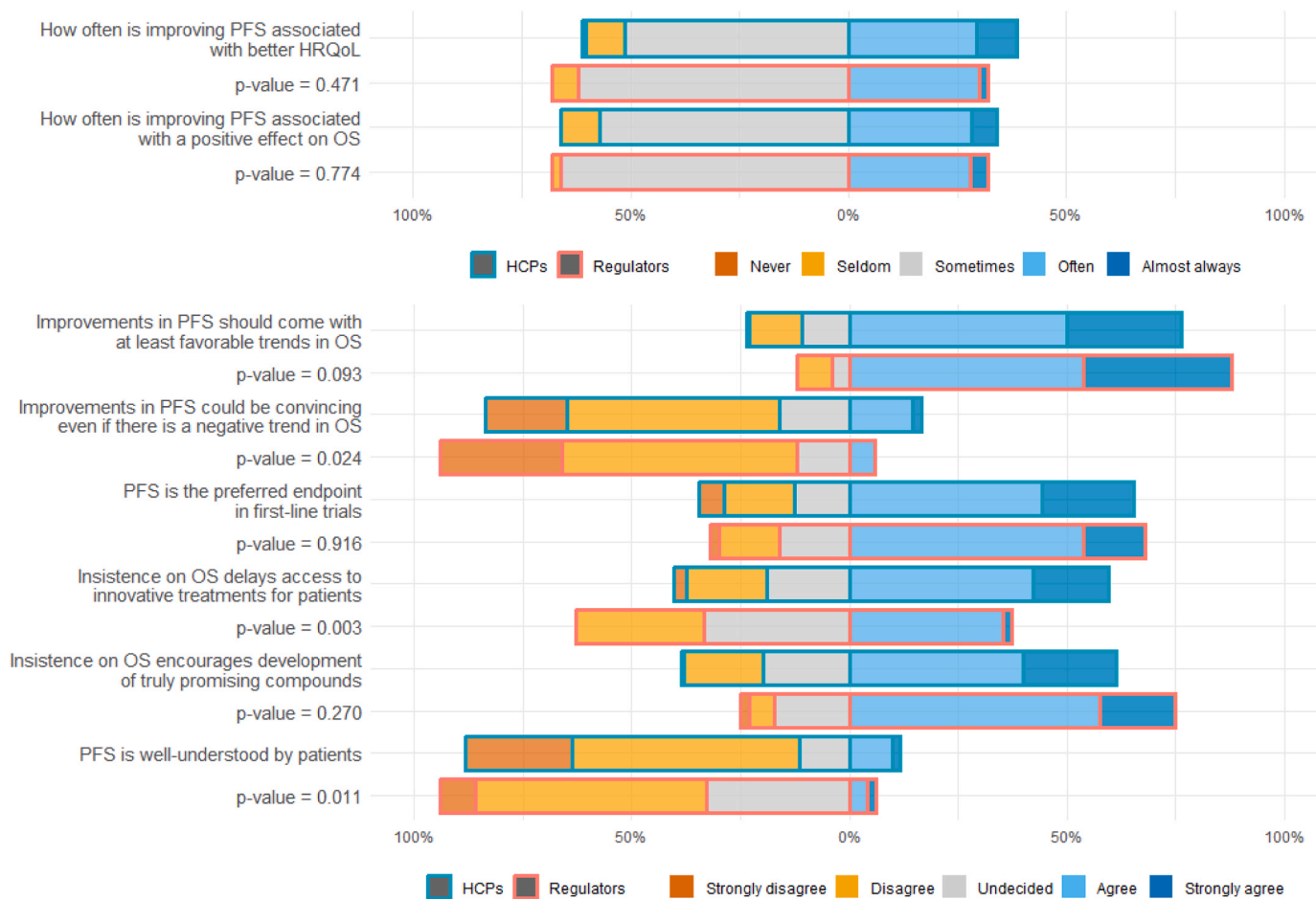
Fig. 2. Responses to the relationship between PFS and other endpoints questions.

Mean scores for the first two factors (PFS as surrogate or intrinsic benefit) were generally higher for participants (and especially for HCPs for factor 2) who declared being less familiar (not at all, slightly, or moderately familiar) with assessing the balance of benefits and harms of cancer drugs, compared to those very or extremely familiar (see Supplementary fig. 2–4).

3.4. Trade-offs between PFS and toxicity

The proportion of HCPs and regulators falling in the different MAR categories for a 6 month improvement in PFS are summarised in Fig. 4. The MAR distributions were spread out in both groups, with a shift towards higher MAR values for the regulators-group (Mann–Whitney U test p-value < 0.001). The MAR distributions did not differ significantly between scenarios with or without supportive OS trend (Mann–Whitney U test p-value = 0.423; Supplementary fig. 5).

The MAR distributions across subgroups defined in terms of tertiles of the estimated scores on the three factors from the exploratory factor analysis are summarised in Supplementary fig. 6–8. Higher estimated scores on factors 1 and 2 were associated with higher median MAR values, with Kruskal–Wallis test p-values of 0.021 and 0.007, respectively. No association was found between the median MAR and the estimated scores on factor 3 (Kruskal–Wallis test p-value = 0.911). Lower familiarity with benefit-risk assessment of cancer drugs was also associated with higher MAR of toxicity for a PFS gain (Kruskal–Wallis test p-value = 0.040; see Supplementary fig. 9).

4. Discussion

This online, cross-sectional survey elicited attitudes about PFS and statements about willingness to trade between PFS and severe toxicity from experienced HCPs and drug regulators. Results need to be taken with caution due to the low number of participants.

Willingness to trade toxicity for gains in PFS was confirmed to be spread within each group of HCPs and regulators. Views were varied also about the relation between PFS and OS or HRQoL, with a majority leaning towards infrequent surrogacy. Regulators showed a slightly greater toxicity risk acceptance compared to HCPs, which is consistent with a previous stated preference study [6]. This may be explained by regulators' allowing a wider range of choices and relying on doctors and patients for benefit-risk assessment in the local context [7]. The large majority of participants' considering that PFS is not well-understood by patients points to the importance of effective communication in clinical decisions.

Responses to the attitudes about PFS questions could be explained in terms of three underlying factors, possibly interpretable as: (i) beliefs about surrogacy for OS and HRQoL, (ii) beliefs about PFS as intrinsic benefit to be distinguished from OS benefit, and (iii) beliefs about the importance of OS. Positive views about the first two factors were associated with a greater willingness to trade toxicity for gains in PFS, attesting to the likely practical implication of such views.

Beliefs about PFS being a surrogate for OS and HRQoL should be further investigated, given the limited situations in which a predictive effect has been established. The importance given to PFS as intrinsic benefit to be distinguished from OS should also be further investigated. Higher scores with respect to this factor were associated with beliefs
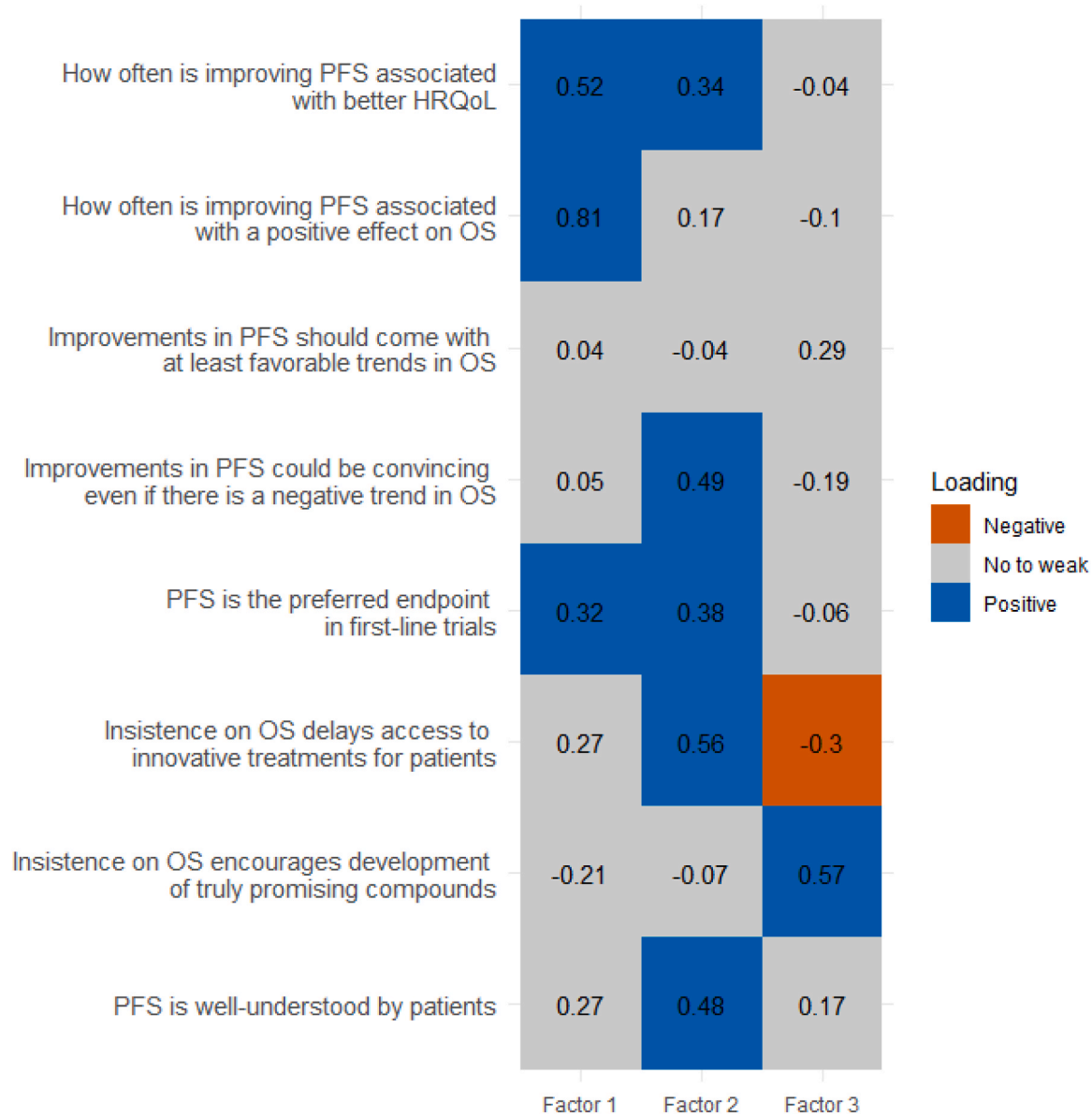
Fig. 3. Estimated factor loadings from the exploratory factor analysis performed on the individual responses to the relationship between PFS and other efficacy endpoints questions. Factor loadings are described as coefficients representing the strength and direction of the relationship between responses (listed on the left of the heatmap) and the three identified factors. The colours in the heatmap visualise how much each observed variable contributes to each factor.

about a positive association between PFS and HRQoL, insistence on OS delaying access to promising treatments, as well as more infrequent responses like believing that PFS is well-understood by patients and that PFS benefits can be important enough to accept a small risk (negative trend) in OS. For both factors, expertise may play a role since participants that declared being less familiar with benefit-risk assessment tended to score higher compared to those that declared being very or extremely familiar.

There was a clear majority in favour of requiring corroborating trends in OS, when PFS is the primary endpoint. Surprisingly, no associations were found between the willingness to trade and the presence for a corroborating trend in OS or the estimated scores on factor 3 (importance of OS). Thus, in practice, the importance of an OS trend may be less than expected.

Despite some illustrative examples [8], bias in the assessment of PFS may have little impact on conclusions in practice, especially when observed effects are large [9]. Still, between one third and one half of respondents in each group were concerned about bias or insufficient

attention given to informative censoring, with strong agreement on the importance of preventing bias with strict rules in the protocol coupled with detailed description of cases of early withdrawal from the study. The likely relative inexperience with statistical methodology among mainly clinical respondents may have played a role. Concerns about bias were not associated with the participants' willingness to trade.

In summary, this survey confirmed varied attitudes of HCPs and regulators about PFS as efficacy endpoint in the advanced cancer setting. The differences related to beliefs about: PFS being a likely surrogate for OS and HRQoL; PFS being an intrinsic benefit to be distinguished from OS; and the importance of OS. Lack of familiarity with benefit-risk assessment was associated with stronger beliefs of PFS being a likely surrogate or an intrinsic benefit. Being a regulator or holding stronger beliefs about PFS being a likely surrogate or an intrinsic benefit were associated with a higher acceptance of risk of toxicity for a PFS gain. Presence of a supportive trend in OS was not associated with a higher acceptable risk. There was agreement on the need to address bias in the adjudication of PFS and the need for improving communication to
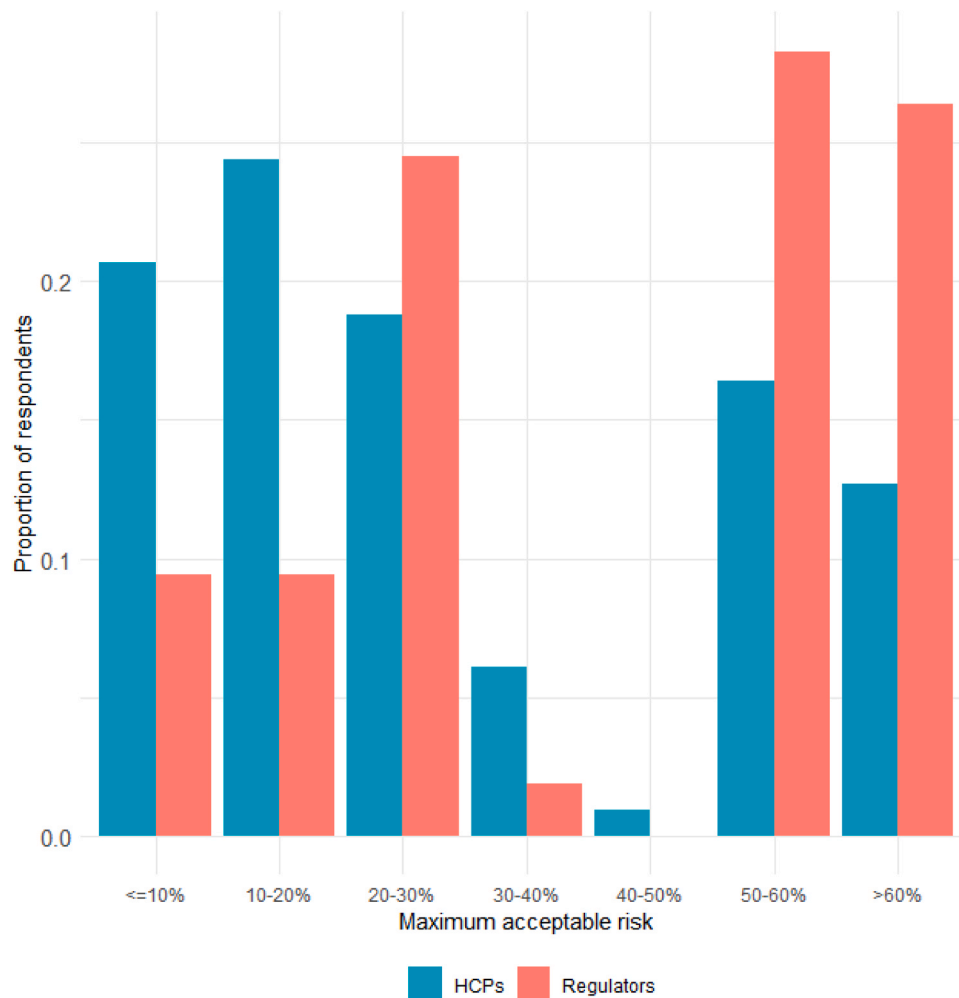
Fig. 4. Comparison of the MAR distributions between HCPs and regulators for a 6 month improvement in median PFS The histogram displays the proportion of HCPs and regulators falling into each maximum acceptable risk category. Each category represents the increase in percent of G34 toxicity risk that is considered maximally acceptable for a 6 month improvement in median PFS.

patients about meaning, strengths, and limitations of improvements in PFS.

Disclaimer

The views expressed in this article are the personal views of the author(s) and may not be understood or quoted as being made on behalf of or reflecting the position of their respective organisations.

Funding

None.

CRediT authorship contribution statement

Douwe Postmus: Conceptualization, Formal analysis, Methodology, Software, Writing – original draft. **Saskia Litiere:** Conceptualization, Writing – review & editing. **Jan Bogaerts:** Conceptualization, Writing – review & editing. **Jurjen Versluis:** Conceptualization, Writing – review & editing. **Jan J. Cornelissen:** Conceptualization, Writing – review & editing. **Francesco Pignatti:** Conceptualization, Investigation, Writing – original draft.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ejca.2023.113496.

References

[1] Fallowfield LJ, Fleissig A. The value of progression-free survival to patients with advanced-stage cancer. Nat Rev Clin Oncol 2012;9(1):41–7. https://doi.org/ 10.1038/nrclinonc.2011.156.

[2] Booth CM, Eisenhauer EA. Progression-free survival: meaningful or simply measurable? J Clin Oncol 2012;30(10):1030–3.

[3] Korn RL, Crowley JJ. Overview: progression-free survival as an endpoint in clinical trials with solid tumors. Clin Cancer Res 2013;19(10):2607–12. https://doi.org/ 10.1158/1078-0432.CCR-12-2934.

[4] Gyawali B, Eisenhauer E, Tregear M, Booth CM. Progression-free survival: it is time for a new name. Lancet Oncol 2022;23(3):328–30. https://doi.org/10.1016/S1470-2045(22)00015-8.

[5] Hauber B, Coulter J. Using the threshold technique to elicit patient preferences: an introduction to the method and an overview of existing empirical applications. Appl Health Econ Health Policy 2020;18(1):31–46. https://doi.org/10.1007/s40258-019-00521-3.

[6] Postmus D, Mavris M, Hillege HL, et al. Incorporating patient preferences into drug development and regulatory decision making: results from a quantitative pilot study with cancer patients, carers, and regulators. Clin Pharm Ther 2016;99(5):548–54. https://doi.org/10.1002/cpt.332.

[7] Pignatti F, Wilking U, Postmus D, Wilking N, Delgado J, Bergh J. The value of anticancer drugs — a regulatory view. Nat Rev Clin Oncol 2022;19(3):207–15. https://doi.org/10.1038/s41571-021-00584-z.

[8] Tannock IF, Pond GR, Booth CM. Biased evaluation in cancer drug trials - how use of progression-free survival as the primary end point can mislead. JAMA Oncol 2022;8 (5):679–80. https://doi.org/10.1001/jamaoncol.2021.8206.

[9] Sridhara R, Mandrekar SJ, Dodd LE. Missing data and measurement variability in assessing progression-free survival endpoint in randomized clinical trials. Clin Cancer Res 2013;19(10):2613–20. https://doi.org/10.1158/1078-0432.CCR-12-2938.