Community-driven development of a modified progression-free survival ratio for precision oncology

Andreas Mock [iD],[1,2,3] Christoph E Heilig,[1,3] Simon Kreutzfeldt,[1,3] Daniel Huebschmann,[3,4,5,6] Christoph Heining,[7,8,9] Evelin Schröck,[9,10] Benedikt Brors,[3,11] Albrecht Stenzinger,[3,12] Dirk Jäger,[2,3] Richard Schlenk,[2,13] Hanno Glimm,[7,8,9] Stefan Fröhling,[1,3] Peter Horak [iD],[1,3] DKTK MASTER Network

For numbered affiliations see end of article.

Correspondence to
Dr Peter Horak;
peter.horak@nct-heidelberg.de

Professor Stefan Fröhling;
stefan.froehling@nct-heidelberg.de

ABSTRACT

Objective Measuring the success of molecularly guided therapies is a major challenge in precision oncology trials. A commonly used endpoint is an intra-patient progression-free survival (PFS) ratio, defined as the PFS interval associated with molecularly guided therapy (PFS2) divided by the PFS interval associated with the last prior systemic therapy (PFS1), above 1.3 or, in some studies, above 1.33 or 1.5.

Methods To investigate if the concept of PFS ratios is in agreement with actual response evaluations by physicians, we conducted a survey among members of the MASTER (Molecularly Aided Stratification for Tumor Eradication Research) Programme of the German Cancer Consortium who were asked to classify the success of molecularly guided therapies in 194 patients enrolled in the MOSCATO 01 trial based on PFS1 and PFS2 times.

Results A comparison of classification profiles revealed three distinct clusters of PFS benefit assessments. Only 29% of assessments were consistent with a PFS ratio threshold of 1.3, whereas the remaining 71% of participants applied a different classification scheme that did not rely on the relation between PFS times alone, but also took into account absolute PFS1 intervals. Based on these community-driven insights, we developed a modified PFS ratio that incorporates the influence of absolute PFS1 intervals on the judgement of clinical benefit by physicians. Application of the modified PFS ratio to outcome data from two recent precision oncology trials, MOSCATO 01 and WINTHER, revealed significantly improved concordance with physician-perceived clinical benefit and identified comparable proportions of patients who benefited from molecularly guided therapies.

Conclusions The modified PFS ratio may represent a meaningful clinical endpoint that could aid in the design and interpretation of future precision oncology trials.

INTRODUCTION

Technological and methodological advances in medicine create opportunities to re-evaluate the standard of care and open new ways to differentiate between success and failure of individual therapies. In oncology, these changes are taking place at breathtaking speed, for example, in the form of next-generation sequencing and advances

Key questions

What is already known about this subject?
► The success of molecularly guided cancer therapies can be measured using intra-patient progression-free survival (PFS) ratios, which are increasingly being used in clinical trial design.

What does this study add?
► The current study investigates the real-life assessments of clinical benefits to molecularly guided cancer therapies based on a survey among in 100 physicians involved in cancer care.
► Our findings led to the development of a new clinical endpoint, the modified PFS ratio.

How might this impact on clinical practice?
► The modified PFS ratio may be used as a novel, clinically relevant endpoint for future precision oncology trials.

in targeted and immunotherapies, and while physicians still struggle to implement these new developments into clinical routine, novel technologies are already looming on the horizon.[1]

To assess the efficacy of new drugs, medical oncology relies on statistical methods and evidence-based principles that were developed in the 1980s and have been continuously advanced and refined since that time. Clinical endpoints acceptable for marketing approval are specified in the regulatory guidelines of the US Food and Drug Administration (Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics) as well as the European Medicines Agency (guideline on the evaluation of anticancer medicinal products in man), and although most therapeutic trials adhere to this framework, there is growing awareness among clinicians that mere statistical improvement might not represent meaningful clinical benefit. As a consequence, several auxiliary scales and

scores have been developed to measure clinical benefit,[2] while overall survival in a randomised controlled trial (RCT) remains the most important endpoint in medical oncology.[3] In addition, surrogate endpoints such as objective response rate, time to progression (TTP), progression-free survival (PFS) as well as symptom control nowadays belong to the standard set of tools for clinical trial design.

With the advent of precision oncology, broadly defined as a strategy to stratify cancer treatment based on molecular data generated by, for example, next-generation sequencing,[4] it has become clear that established clinical endpoints might not be sufficient to assess the efficacy of individualised therapies. This is mainly attributable to the unique molecular profiles of individual tumours, which preclude the compilation of adequately sized uniform patient cohorts required for a RCT. In addition, a placebo or standard-of-care arm may be considered unethical when testing a highly effective targeted drug in patients with advanced-stage cancers who have exhausted conventional treatment options. Thus, the established frequentist, non-Bayesian approach to the statistical design and assessment of clinical trials is difficult to reconcile with the principles of precision oncology. This dilemma was first recognised by Von Hoff and colleagues who developed the growth modulation index, or PFS ratio (PFSr), as a surrogate measure of treatment benefit and novel endpoint for precision oncology studies.[5 6] Paired analysis of individual patients through assessment of the TTP or PFS during and after molecularly guided treatment (TTP2 or PFS2) in relation to the TTP or PFS during and after the last prior systemic therapy (TTP1 or PFS1) better accounted for the interindividual heterogeneity of tumours, patients, and responses to specific drugs. This approach, where each patient serves as his/her own control, circumvents the need for a control arm and has been widely adopted in current precision oncology trials.[7–9] Despite some methodological concerns,[10] several fixed PFSrs (1.3, 1.33, 1.5) have been used as thresholds for the assessment of clinical benefit in recent studies.[7 8]

In this study, we have analysed the strengths and weaknesses of the PFSr as an endpoint in precision oncology trials. Specifically, we describe a community-driven effort towards an improved, modified PFSr (mPFSr) that faithfully represents the judgement of therapeutic benefit by clinicians from multiple institutions and validate the model on available clinical data.

MATERIAL AND METHODS
Survey on clinical benefit assessments based on PFS times
To investigate if the concept of PFSr is in line with actual response evaluations by clinicians, we reached out to referring and/or collaborating physicians within the MASTER (Molecularly Aided Stratification for Tumor Eradication Research) Programme of the German Cancer Consortium[11] with a classification task. The survey was based on PFS times associated with molecularly guided therapies within the MOSCATO 01 precision oncology

trial[7] (figure 1A), and participants were asked to categorise patients as 'responder' or 'non-responder' based on PFS1 and PFS2 (figure 1B). We deliberately provided no guidelines for evaluation of the data but explicitly requested the participants' individual response assessment. Alongside the classification, we collected information on the physicians' profiles, including age, years of practice, and regular participation in molecular tumour boards (online supplementary table S1). We received 100 complete replies, which were included in the analysis.

Statistical analyses
Statistical analyses were carried out in R (V.3.5.2) with the extended functionality of the *tidyverse* package (V.1.2.1).[12] The binary heatmap was generated using the *ComplexHeatmap* package (V.1.99.8).[13] The alluvial plot was drawn with the *ggalluvial* package (V.0.9.1). Statistical significance was assessed using t-tests or a Wilcoxon signed-rank test; *p<0.05, **p<0.01, ***p<0.001.

RESULTS
Physician-perceived clinical benefit of molecularly guided therapies
To assess the congruence of response evaluations by physicians, we first calculated the fraction of votes classifying a case as responder. Visualising the distribution of fractions (figure 1C) revealed distinct peaks at 0 (categorisation as responder by 0% of participants) and 1 (categorisation as responder by 100% of participants), indicating general agreement among physicians in the perceived clinical benefit of molecularly guided therapies. However, approximately 25% of cases were categorised as responder or non-responder by a similar number of survey participants.

Cluster analysis revealed three distinct groups of classifications (figure 1D). Interestingly, only 29% of classifications were consistent with a PFSr threshold of 1.3 (Cluster 3), whereas the remaining 71% of physicians applied a significantly different classification scheme (p<0.001, figure 1E). For example, Cluster 1 was characterised by a significantly higher fraction of responders (p<0.001, figure 1F). Neither age or years of practice of the survey participants were significantly associated with membership to a specific cluster. However, physicians that only occasionally participate in molecular tumour boards were enriched in Cluster 2 (p=0.039, online supplementary table S2).

Discordance between physician-perceived clinical benefit and PFSr threshold
Next, we systematically analysed the discordance between physician-perceived clinical benefit, which was considered as the ground truth, and a PFSr threshold of 1.3. As expected, the physician classifications were significantly associated with PFS2 duration (r=0.79, p<2.2E-16); however, there was a discordance between the majority vote of the physicians (responder class fraction >0.5) and the PFSr threshold (figure 2A) in 16% of cases
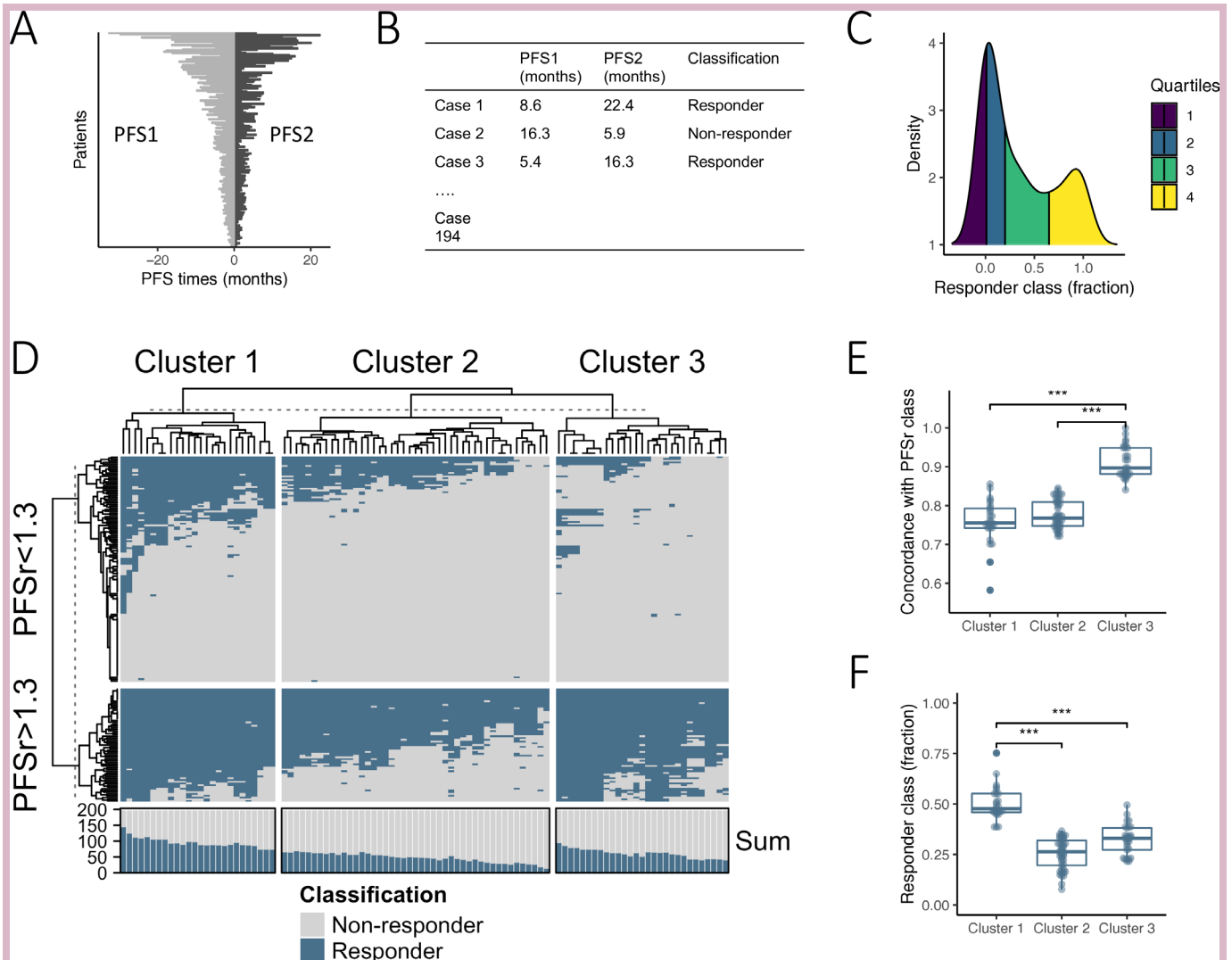
Figure 1 Physician-perceived clinical benefit of molecularly guided therapy. (A) PFS1 and PFS2 times of 194 patients enrolled in the MOSCATO 01 trial. (B) Examples of response assessments by physicians based on PFS1 and PFS2 times. (C) Distribution of fraction of responder class assignments. For example, a fraction of 1 denotes that all physicians classified the respective case as responder to molecularly guided therapy. (D) Binary heatmap of response classifications (rows) by 100 physicians (columns) showing three distinct clusters of assessments (k-means clustering with k=3). Classifications were stratified by PFSr threshold. Stacked bar plots indicate the sum of response classifications (bottom). (E) Boxplots comparing the row-wise fraction of responder class assignments by physicians between the three clusters. (F) Boxplots comparing the concordance of response classifications by physicians with the class defined by a PFSr above 1.3 between the three clusters. *p<0.05, **p<0.01, ***p<0.001. PFS, progression-free survival; PFSr, PFS ratio.

(false positive, 9%; false negative, 7%; figure 2B,C). To understand these discrepancies, we looked into associations between incorrect classifications and PFS duration and observed that false positive cases were attributable to very short PFS1 times of less than 2 months, whereas false negative cases could be explained by similarly long PFS1 and PFS2 times resulting in low PFSr values (figure 2D).

Community-driven definition of a modified PFSr

The analysis of response classifications by physicians revealed two major limitations of the current PFSr. On one hand, the clinical benefit of molecularly guided therapies will be overestimated if the PFS1 duration is very short. For example, a patient whose PFS1 and PFS2 times are 1 and 2 months, respectively, will be classified as deriving clinical benefit based on a PFSr of 2. On the other hand, the efficacy of a molecularly guided therapy might be undervalued if the PFS1 and PFS2 intervals are identical. For example, a patient whose PFS1 and PFS2 times are both 12 months will be classified as non-responder based on a PFSr of 1. To overcome these limitations, we propose a modified PFSr (mPFSr; figure 3A, equation 1) that comprises two simple modifications to the conventional PFSr. First, to adjust for false positives, all PFS1 times below 2 months are converted into a modified PFS1 interval, termed pre-PFS, of 2 months (figure 3A, equation 2). Second, to correct for false negative results, PFS2
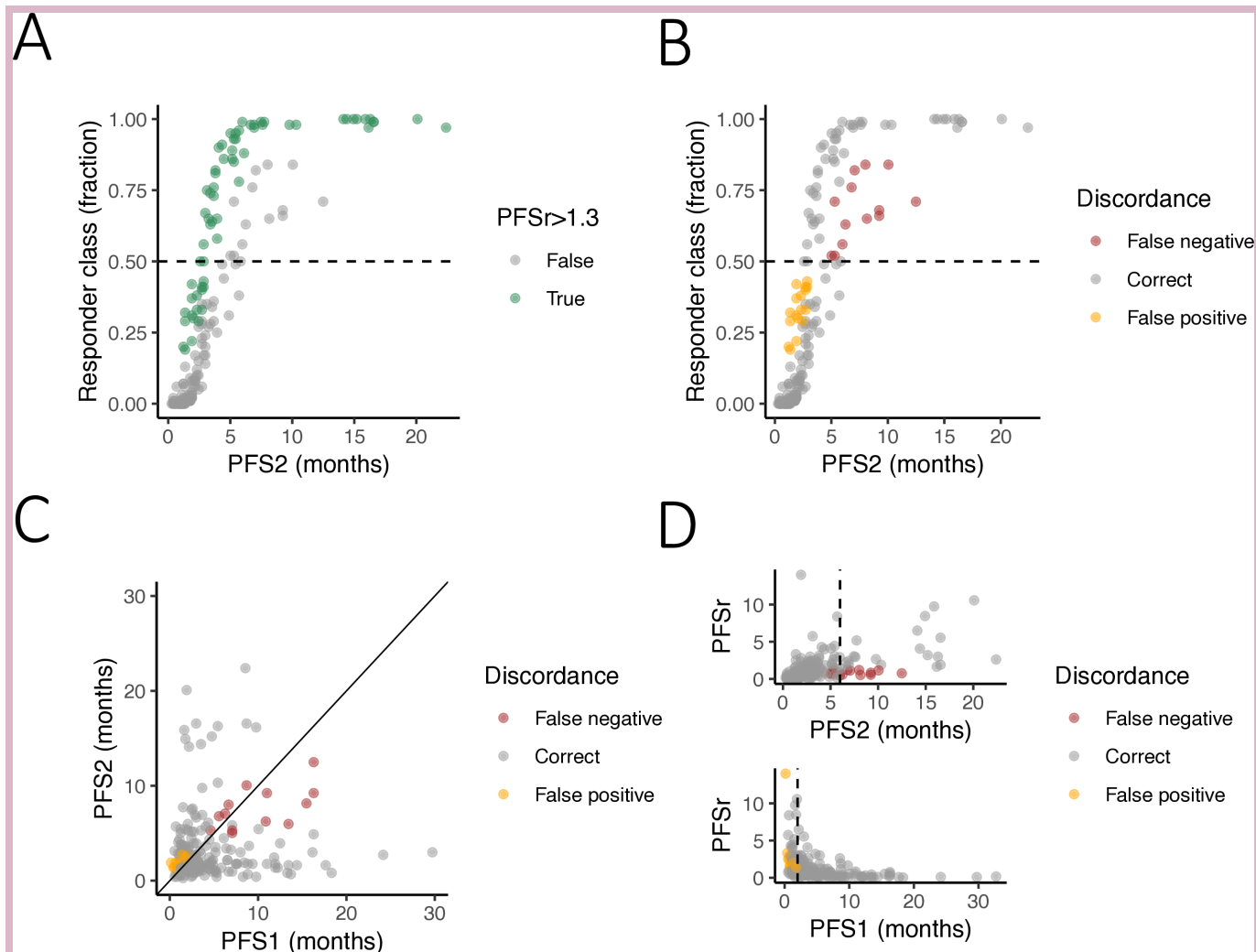
Figure 2 Discordance between physician-perceived clinical benefit and PFSr threshold. (A) Fraction of cases classified as responders plotted over PFS2 time, coloured according to PFSr. (B) Fraction of cases classified as responders plotted over PFS2 time, coloured according to discordance between physician classification, which was considered as the ground truth, and PFSr threshold. (C) PFS2 time plotted over PFS1 time, coloured according to discordance between physician classification and PFSr threshold. (D) PFSr plotted over PFS2 (top) and PFS1 (bottom), coloured according to discordance between physician classification and PFSr threshold. PFS, progression-free survival; PFSr, PFS ratio.

times above 6 months are assumed to reflect successful molecular guided therapy and are set to 24 months (post-PFS), resulting in PFSr above the currently used thresholds of 1.3, 1.33, or 1.5.

Comparative analysis of PFSr and mPFSr
To investigate whether the mPFSr more accurately reflects the clinical benefit of molecularly guided therapies, we compared the survey results with the categorisation of patients based on a PFSr or mPFSr threshold of 1.3 and observed that the concordance with the physician classification was significantly higher for the mPFSr (figure 3B). Furthermore, application of the mPFSr to the outcome data of two recent precision oncology trials, MOSCATO 01 and WINTHER,[7 8] identified comparable proportions of patients who benefited from molecularly guided therapies (33% vs 31% and 24% vs 24%, respectively; figure 3C,D).

DISCUSSION
In recent years, several precision oncology trials have used a PFSr as the primary endpoint. For example, the MOSCATO 01 study applied a PFSr threshold of 1.3 to detect a clinical benefit of molecularly targeted therapies in 33% of patients.[7] The negative SHIVA study, on the other hand, the first randomised trial of molecularly targeted cancer therapy versus treatment of physician's choice, was one of the few precision oncology trials that did not use a PFSr but a standard PFS endpoint.[14] Nevertheless, this study highlighted important challenges of precision oncology in identifying histology-independent oncogenic drivers, in delivering clinically effective targeted therapies, and in adequate trial design.[15] All the more interesting, a secondary, retrospective cross-over analysis of the SHIVA data revealed that 37% of patients who received molecularly targeted therapy after progression on prior systemic
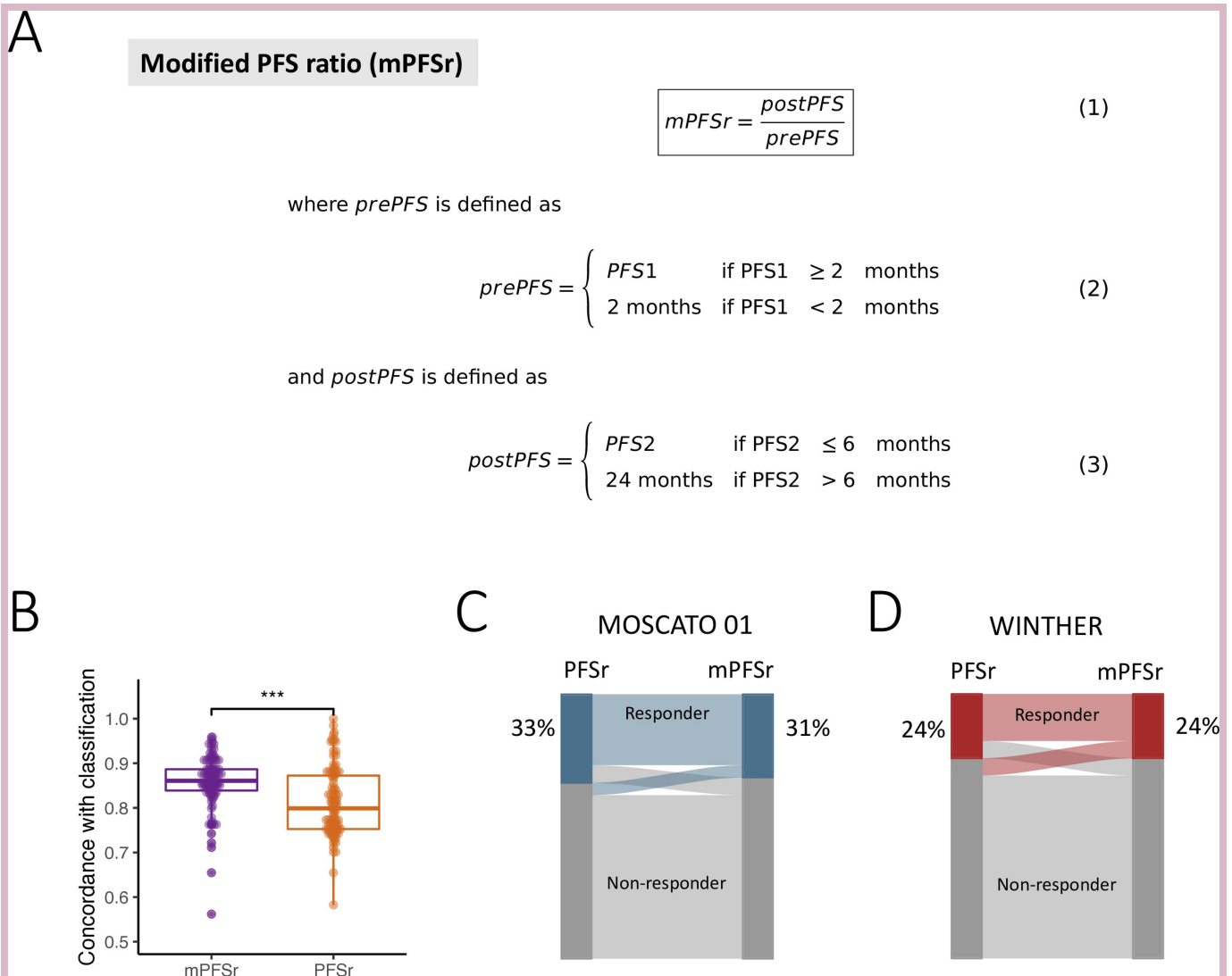
Figure 3 Definition of a mPFSr. (A) Algebraic definition of a mPFSr. In contrast to standard PFSr, the PFS2 interval is divided by the so-called prePFS time to correct for false positive predictions. In addition, PFS2 times above 6 months are set to 24 months to correct for false negatives (postPFS). (B) Boxplots comparing the concordance with the physician classification between PFSr and mPFSr. *p<0.05, **p<0.01, ***p<0.001. (C, D) Alluvial diagrams comparing the classification of patients from the MOSCATO 01 (C) and WINTHER (D) trials according to PFSr and mPFSr thresholds of 1.3. mPFSr, modified PFSr; PFS, progression-free survival; PFSr, PFS ratio.

treatment had a PFSr above 1.3.[16] The recently published WINTHER trial aimed to achieve a PFSr of more than 1.5 in at least 50% of patients (40% for therapies based on RNA sequencing data). While this ambitious primary endpoint was not met, WINTHER demonstrated for the first time the feasibility of international prospective precision oncology trials.[8] Despite its increasing popularity, the PFSr has been systematically investigated only in a handful of studies.[10 17–20]

We aimed to understand how the PFSr relates to physician-perceived clinical benefit, as judged based on PFS1 and PFS2 times alone. In addition to several clusters of PFS benefit evaluations, we identified PFS2/PFS1 relationships that likely result in false positive or false negative classifications when applying a fixed PFSr threshold. Specifically, most false positive assessments were caused by short PFS1 times of less than 2 months, whereas false negative assignments could be attributed to PFS2 and PFS1 intervals of equal length. On the basis of these community-driven insights, we developed a mPFSr that better reflects the clinical benefit assessment of physicians involved in cancer care, and reanalysis of outcome data from published precision oncology studies[7 8] using the mPFSr revealed a comparable proportion of patients who benefited from molecularly guided therapies.

In summary, we believe that the mPFSr represents a meaningful clinical endpoint that will aid in the design and interpretation of future precision oncology trials. There remain a number of important tasks, such as the application of uniform response evaluation criteria, systematic quality-of-life assessments, and the consideration of PFS times prior to PFS1, and we are confident

that current efforts to meet these challenges will yield the next generation of outcome measures.

Author affiliations
[1]Department of Translational Medical Oncology, National Center for Tumor Diseases (NCT) Heidelberg and German Cancer Research Center (DKFZ), Heidelberg, Germany
[2]Department of Medical Oncology, NCT Heidelberg and Heidelberg University Hospital, Heidelberg, Germany
[3]German Cancer Consortium (DKTK), Heidelberg, Germany
[4]Computational Oncology, Molecular Diagnostics Program, NCT Heidelberg and Heidelberg University Hospital, Heidelberg, Germany
[5]Heidelberg Institute for Stem Cell Technology and Experimental Medicine (HI-STEM gGmbH), Heidelberg, Germany
[6]Department of Pediatric Immunology, Hematology and Oncology, Heidelberg University Hospital, Heidelberg, Germany
[7]University Hospital Carl Gustav Carus, Dresden, Germany
[8]Department of Translational Medical Oncology, NCT Dresden, Dresden, Germany
[9]DKTK, Dresden, Germany
[10]Institute for Clinical Genetics, Faculty of Medicine Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany
[11]Division of Applied Bioinformatics, DKFZ and NCT Heidelberg, Heidelberg, Germany
[12]Institute of Pathology, Heidelberg University Hospital, Heidelberg, Germany
[13]NCT Trial Center, NCT and DKFZ Heidelberg, Heidelberg, Germany

Collaborators DKTK Master Network (alphabetical order): Leonidas Apostolidis, Marinela Augustin, Daniela Aust, Irfan Ahmed Bhatti, Johannes Bloehdorn, Cornelia Brendel, Christian Britschgi, Jan Braess, Stefan Burdach, Elena Busch, Jozefina Casuscelli, Alexander Desuki, Thomas Deutsch, Mareike Dietrich, Ursula Ehmer, Thomas J. Ettrich, Johanna Falkenhorst, Tanja Fehm, Anne Flörcken, Andrea Forschner, Stefan Fuxius, Maria Gonzales-Carmona, Frank Griesinger, Sabine Grill, Stefan Gröschel, Georg Martin Haag, Ulrich Haag, Niels Halama, Holger Hebart, Nina Heidger, Barbara Hermes, Georg Hess, Simone Hettmer, Manuela Hoechstetter, Martin Hoffmann, Felix J. Hüttner, Anna L. Illert, Maximilian Jenzer, Bernd Kasper, Stefan Kasper-Virchow, Thomas Kindler, Ewa Koscielniak, Jan Krönke, Michael Kühn, Volker Kunzmann, Alois Lang, Jonas Leichsenring, Elisabeth Livingstone, Lucia Liotta, Kim Luley, Elisabeth Mack, Uwe M. Martens, Klaus Metzeler, Jan Moritz Middeke, Lino Möhrmann, Roopa Jayarama-Naidu, Ulrich-Frank Pape, Lukas Perkhofer, Arne Pfeufer, Constantin Pixberg, Michael Quante, Bernhard Rendenbach, Damian Rieke, Christian Rothermundt, Andre Norbert Sagerer, Martin Salzmann, Dieter Saur, Bastian Schilling, Jan Schleicher, Anke Schlenska-Lange, Thomas Schmidt, Sophia Schmitz, Sebastian Schölch, Rajiv Shah, Khalid Shoumariyeh, Alexander Siebenhüner, Martin Singh, Jens Siveke, Christoph Springfeld, Helen Starke, Sophia Strobel, Veronica Teleanu, Niklas Thon, Sebastian Wagner, Thomas Walle, Benedikt Westphalen, Bettina Whitlock, Eva Winkler, Naita Maren Wirsik, Lena Woydack, Angelika Zabel-du Bois, Stefanie Zschäbitz. More details about collaborators are presented in the online supplementary material.

Contributors Conception of the work: AM, SF and PH. Funding acquisition: AM, SF, DJ. Data collection and data analysis: AM, CEH, SK, DH. Manuscript writing/editing: AM, CEH, SK, DH, AS, SF. Critical revision of the manuscript: CH, ES, BB, DJ, AS, SF, RS, HG. Final approval: all authors. Collaborators of the DKTK MASTER programme have answered the survey. Shared last authorship: HG, SF, PH.

Competing interests SF has had a consulting or advisory role, received honoraria, research funding, and/or travel/accommodation expenses from the following for-profit companies: Bayer, Roche, Amgen, Eli Lilly, PharmaMar, AstraZeneca, and Pfizer.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information.

ORCID iDs
Andreas Mock http://orcid.org/0000-0002-3332-9166
Peter Horak http://orcid.org/0000-0003-4536-9306

REFERENCES

1 Subbiah V, Kurzrock R. Challenging standard-of-care paradigms in the precision oncology era. *Trends Cancer* 2018;4:101–9.
2 Cherny NI, de Vries EGE, Dafni U, *et al*. Comparative assessment of clinical benefit using the ESMO-magnitude of clinical benefit scale version 1.1 and the ASCO value framework net health benefit score. *J Clin Oncol* 2019;37:336–49.
3 Driscoll JJ, Rixe O. Overall survival: still the gold standard. *Cancer J* 2009;15:401–5.
4 Moscow JA, Fojo T, Schilsky RL. The evidence framework for precision cancer medicine. *Nat Rev Clin Oncol* 2018;15:183–92.
5 Von Hoff DD, Stephenson JJ, Rosen P, *et al*. Pilot study using molecular profiling of patients' tumors to find potential targets and select treatments for their refractory cancers. *J Clin Oncol* 2010;28:4877–83.
6. Von Hoff DD. There are no bad anticancer agents, only bad clinical trial designs-twenty-first Richard and Hinda Rosenthal Foundation Award Lecture. *Clin Cancer Res* 1998.
7 Massard C, Michiels S, Ferté C, *et al*. High-Throughput genomics and clinical outcome in Hard-to-Treat advanced cancers: results of the Moscato 01 trial. *Cancer Discov* 2017;7:586–95.
8 Rodon J, Soria J-C, Berger R, *et al*. Genomic and transcriptomic profiling expands precision cancer medicine: the WINTHER trial. *Nat Med* 2019;25:751–8.
9 Sicklick JK, Kato S, Okamura R, *et al*. Molecular profiling of cancer patients enables personalized combination therapy: the I-PREDICT study. *Nat Med* 2019;25:744–50.
10 Wu J, Chen L, Wei J, *et al*. Phase II trial design with growth modulation index as the primary endpoint. *Pharm Stat* 2019;18:212–22.
11 Horak P, Klink B, Heining C, *et al*. Precision oncology based on omics data: the NCT Heidelberg experience. *Int J Cancer* 2017;141:877–86.
12 Wickham H. Tidyverse: Easily Install and Load "Tidyverse" Packages 2016.
13 Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 2016;32:2847–9.
14 Le Tourneau C, Delord J-P, Gonçalves A, *et al*. Molecularly targeted therapy based on tumour molecular profiling versus conventional therapy for advanced cancer (SHIVA): a multicentre, open-label, proof-of-concept, randomised, controlled phase 2 trial. *Lancet Oncol* 2015;16:1324–34.
15 Le Tourneau C, Kurzrock R. What have we learned from SHIVA? *Nat Rev Clin Oncol* 2016;13:719–20.
16 Belin L, Kamal M, Mauborgne C, *et al*. Randomized phase II trial comparing molecularly targeted therapy based on tumor molecular profiling versus conventional therapy in patients with refractory cancer: cross-over analysis from the SHIVA trial. *Ann Oncol* 2016.
17. Texier M, Rotolo F, Ducreux M, *et al*. Evaluation of treatment effect with paired failure times in a single-arm phase II trial in oncology. *Comput Math Methods Med* 2018;2018:1–8.
18 Cousin S, Taieb S, Penel N. A paradigm shift in tumour response evaluation of targeted therapy: the assessment of novel drugs in exploratory clinical trials. *Curr Opin Oncol* 2012;24:338–44.
19 Buyse M, Quinaux E, Hendlisz A, *et al*. Progression-Free survival ratio as end point for phase II trials in advanced solid tumors. *J Clin Oncol* 2011;29:e451–2.
20 Kovalchik S, Mietlowski W. Statistical methods for a phase II oncology trial with a growth modulation index (GMI) endpoint. *Contemp Clin Trials* 2011;32:99–107.