PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Safety and efficacy of regorafenib in patients with treatment-refractory metastatic colorectal cancer in Turkey: the single-arm, open-label REGARD study
AUTHORS	Dane, Faysal; Ozgurdal, Kirhan; Yalçın, Şuayib; Benekli, Mustafa; Aykan, Nuri Faruk; Yücel, İdris; Özkan, Metin; Evrensel, Turkkan; Sevinç, Alper; Coskun, Hasan Şenol; Sanli, Ulus Ali; Kara, Ismail Oguz; Yumuk, Perran Fulden

VERSION 1 – REVIEW

REVIEWER	Chaoyuan Kuang, M.D., Ph.D. Medical Oncology Fellow UPMC Hillman Cancer Center Pittsburgh, PA, United States
REVIEW RETURNED	13-Feb-2019

GENERAL COMMENTS	Overall my biggest criticism of this manuscript is the claim that a phase III trial was conducted. The study itself claims that he primary objective was safety and there is no control group in this trial to compare efficacy to. There is also no mention of the most recent studies and abstracts regarding regorafenib use in colorectal cancer, or any attempt at performing a statistically rigorous historical control. I have comments elaborating on my checklist points below. 1. Checklist item 3: The study design is for safety as the primary objective. In my opinion a study with this aim should not be designated a phase III study. Furthermore the efficacy analysis is somewhat minimalist. This could be greatly improved if a strong historical control cohort were generated and outcomes comparison was performed. Along these lines, while I do not suspect any redundancy of publication, the CORRECT trial did include study sites in Turkey. While the patient number here will likely be small, it may be worthwhile to investigate if the results in these Turkish CORRECT trial patients roughly correlate to the findings here. 2. Checklist item 8: I would recommend incorporation of the latest results from the Redose trial (study of regorafenib at multiple doses) into the discussion section. This study has some intriguing early results suggesting that lower dose regorafenib as a starting dose with the ability to escalate dose based on tolerability, may lead to improved overall survival. 3. Checklist item 12: If the authors chose to maintain this analysis as is, without further analysis or historical control cohort, then I think it would be appropriate to acknowledge the lack of control

	and limitations this places on any efficacy analysis. For example since the larger phase III studies cited here are multinational studies, could there be some difference in best supportive care (the appropriate control arm for last line therapy) in Turkey, that either makes regorafenib more, or less efficacious compared to the the average international CRC patient? This is worthy of exploration, and ideally a true phase III placebo controlled trial would be run to answer for Turkish patients if this medicine is efficacious.

REVIEWER	Nobumichi Takeuchi Department of medical oncology Ina Central hospital
REVIEW RETURNED	27-Mar-2019

GENERAL COMMENTS	Regorafenib requires tactile management of AE to use it for long time. At this point, this report has value to publish, however REDOSE study presented by Bekkai-Saab on ASCO 2018 seemed to be needed to mention.

REVIEWER	Shouki Bazarbashi King Faisal Specialist Hospital and Research center Riyadh, Saudi Arabia
REVIEW RETURNED	24-May-2019

GENERAL COMMENTS	Well designed and written study. Answered the addressed question which is the toxicity and efficacy of regorafenib in Turkish patients. There are several minor points that need to be addressed by the author: 1. the author should correct the design of the study, This is a phase II or phase IIIb rather than phase III study. This should be corrected in the manuscript and abstract 2. since disease progression was either clinical or radiological, the authors should report the percentage and number of patients whose disease progression was on clinical ground. 3. since the timing of the evaluation of response was according to individual institutional policy the author should report time to 1st radiological evaluation. 4. Page 8, the last sentence should read: all patients have received prior treatment with fluoropyrimidine analogs, oxaliplatin, Irinotecan and monoclonal antibodies (Bevacizumab, cetuximab/panitumumab) 5. table 1 reports that 77 patients did not have brain metastasis. The author should clarify if CT brain was one of the pre-study evaluation and whether it was optional? 6. page 10, under safety, the author should define the difference between grade 3-4 treatment-emergent adverse events (TEAE) and serious TEAE 7. page 11, table 2, worst grade AE. The number of drug-related treatment emergent grade 1 and 2 AE are more than treatment-related AE. This does not make sense. Ther drug-related should be less, not more. The author should review those figures and correct them.

REVIEWER	Mike Bradburn University of Sheffield. UK
REVIEW RETURNED	15-Sep-2019

GENERAL COMMENTS	The study reports a single arm clinical evaluation of Regorafenib for the treatment of refractory metastatic colorectal cancer conducted in 11 Turkish cancer hospitals. Previous randomised evidence for Regorafenib (references 5&6) concluded Regorafenib has a modest but statistically significant improvement on overall survival, tempered by an increased number of grade 3 or 4 toxic adverse events. The study was intended as a descriptive, uncontrolled summary of Regorafenib in the Turkish population, primarily to quantify adverse events but also progression free survival. The outcomes are consistent with those reported on the previous randomised Regorafenib. The report appears an accurate and balanced description of the study and its outcomes are consistent with those planned in its clinicaltrials.gov record (NCT01853319). As an uncontrolled trial the most relevant reporting standards are a mixture of CONSORT and STROBE, but these appear have been met with one exception discussed below. The comments below are minor, but my main question concerns the importance of this to the journal readership. The focus of this is of limited interest to a international, general medical journal: it is hard to know what this adds. My own background (medical statistician) means I am unable to answer this definitively but I believe the new information presented in this trial is of very limited interest. At the least, some justification for why the two international trial findings (plus other referenced studies) do not address safety & efficacy for the Turkish population - or in other words, explanation on why the Turkish population is in some sense different - would help. I have two more specific comments: [Major] 1) No sample size justification is presented. The number of patients (100) does seem a reasonable choice for this design and could be retrospectively justificed on grounds of precision (standard error) but the manuscript does not meet generic reporting guidance. [Minor] 2) The two previous randomised trials assessed progression using RECIST criteria whereas this has used a more loose definition. Again this may be reasonable since progression free survival is a secondary interest, but should be noted more clearly as a limitation.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1	General	Overall my biggest criticism of this manuscript is the claim that a phase III trial was conducted. The study itself claims	As stated in the response to the Associate Editor above (Associate Editor, item 1), REGARD was registered as a phase III study because the Turkish health regulations did not have a provision for phase IIIb studies. It is described as a phase III study on the Bayer website	

		that the primary objective was safety and there is no control group in this trial to compare efficacy to. There is also no mention of the most recent studies and abstracts regarding regorafenib use in colorectal cancer, or any attempt at performing a statistically rigorous historical control. I have comments elaborating on my checklist points below.	(https://clinicaltrials.bayer.com/study/2023) and on ClinicalTrials.gov (https://clinicaltrials.gov/ct2/show/NCT01853319). However, the design of REGARD is more accurately described as a phase IIIb study, so we have revised the descriptions in the manuscript to “phase IIIb”. The primary objectives of REGARD were to evaluate the safety of regorafenib in patients with mCRC and to estimate progression-free survival (PFS). This study design is similar to that of the large, recently published CONSIGN study (Van Cutsem, et al. Oncologist 2019;24:185–192). Although there was no control group, the efficacy results (PFS) can be put in the context of other studies of regorafenib for mCRC. This is currently done in the Discussion section, referencing the CORRECT, CONCUR, and CONSIGN trials. In addition, we have incorporated results from the more recent ReDOS and CORRELATE studies into the Discussion section. More details are provided below.	3, 5, 7 18, 19
Reviewer 1	1a	Checklist item 3: The study design is for safety as the primary objective. In my opinion a study with this aim should not be designated a phase III study.	As stated above (Reviewer 1 general and Associate Editor item 1), we have now revised the description of the study design to “phase IIIb”.	3, 5, 7
Reviewer 1	1b	Furthermore the efficacy analysis is somewhat minimalist. This could be greatly improved if a strong historical control cohort were generated and outcomes comparison was performed.	Regorafenib was the first tyrosine kinase inhibitor approved for the treatment of mCRC in the third line and beyond, so historical controls are not available. The reviewer is correct that the efficacy analysis provides limited information. As stated in the Methods (p. 8) and Discussion (p. 18) sections, PFS was based on investigator assessment, with the frequency of assessments based on each institution’s best standard of care. The lack of standardized PFS assessments across study sites makes it difficult to compare the efficacy results with those from randomized controlled trials. However, the large, phase IIIb CONSIGN study (N=2864) had a similar design to that of	8, 18, 19

			REGARD, and was conducted at approximately the same time across 25 countries in Europe, North America, Israel, and Australia (Van Cutsem, et al. Oncologist 2019;24:185–192), and the efficacy results of CONSIGN are similar to those found in REGARD. In the Discussion section, we have added text and a reference to help put the results from this study in the context of existing data from the larger, CONSIGN study. In addition, we added a reference to the large, observational CORRELATE study that was carried out in >1000 patients and reported a similar PFS (Ducreux M, et al. Eur J Cancer 2019, 123:146-54).	
Reviewer 1	1c	Along these lines, while I do not suspect any redundancy of publication, the CORRECT trial did include study sites in Turkey. While the patient number here will likely be small, it may be worthwhile to investigate if the results in these Turkish CORRECT trial patients roughly correlate to the findings here.	Although there were study sites for the CORRECT trial in Turkey, no patients from Turkey were randomized in CORRECT. We have added text to the Introduction and Discussion sections to clarify that no patients from Turkey were randomized or treated in CORRECT, CONCUR, or CONSIGN.	6, 17
Reviewer 1	2	Checklist item 8: I would recommend incorporation of the latest results from the Redose trial (study of regorafenib at multiple doses) into the discussion section. This study has some intriguing early results suggesting that lower dose regorafenib as a starting dose with the ability to escalate dose	A paragraph incorporating the results of the ReDOS trial has been added to the Discussion section (p. 18).	18

		based on tolerability, may lead to improved overall survival.		
Reviewer 1	3	Checklist item 12: If the authors chose to maintain this analysis as is, without further analysis or historical control cohort, then I think it would be appropriate to acknowledge the lack of control and limitations this places on any efficacy analysis. For example since the larger phase III studies cited here are multinational studies, could there be some difference in best supportive care (the appropriate control arm for last line therapy) in Turkey, that either makes regorafenib more, or less efficacious compared to the the average international CRC patient? This is worthy of exploration, and ideally a true phase III placebo controlled trial would be run to answer for Turkish patients if this medicine is efficacious.	REGARD was designed primarily as a safety study with no control arm because before its initiation, the randomized controlled CORRECT trial met its primary endpoint of a statistically significant improvement in overall survival. It would not have been ethical to include a placebo arm in the REGARD study. Text has been added to the Discussion section highlighting and explaining the lack of a control arm, and stating the limitations on interpreting the efficacy results (p 18, 19).	18, 19
Reviewer 2	1	Regorafenib requires tactile management of AE	We agree with the reviewer that managing treatment-related adverse events to allow patients to continue therapy is key. We have	18

		to use it for long time. At this point, this report has value to publish, however REDOSE study presented by Bekkai-Saab on ASCO 2018 seemed to be needed to mention.	added a paragraph to the Discussion section (p. 18) incorporating the results from the ReDOS trial to highlight that using a dose-escalation strategy is a viable option that may allow patients to remain on therapy longer.	
Reviewer 3	General	Well designed and written study. Answered the addressed question which is the toxicity and efficacy of regorafenib in Turkish patients. There are several minor points that need to be addressed by the author:	We thank the reviewer for this comment.	
Reviewer 3	1	The author should correct the design of the study, This is a phase II or phase IIIb rather than phase III study. This should be corrected in the manuscript and abstract.	As stated earlier in our responses to Reviewer 1 (general and item 1a) and the Associate Editor (item 1), we have revised the description of the study design to “phase IIIb”.	
Reviewer 3	2	Since disease progression was either clinical or radiological, the authors should report the percentage and number of patients whose disease progression was on clinical ground.	Only one patient (1%) had disease progression diagnosed clinically. This is shown in Supplemental Figure 1.	
Reviewer 3	3	Since the timing of the evaluation of response was according to individual institutional policy	The time to first radiologic evaluation was not collected in the REGARD trial.	

		the author should report time to 1st radiological evaluation.		
Reviewer 3	4	Page 8, the last sentence should read: all patients have received prior treatment with fluoropyrimidine analogs, oxaliplatin, Irinotecan and monoclonal antibodies (Bevacizumab, cetuximab/panitumumab)	This change has been made.	10
Reviewer 3	5	Table 1 reports that 77 patients did not have brain metastasis. The author should clarify if CT brain was one of the pre-study evaluation and whether it was optional?	A CT scan of the brain was part of the screening evaluation and was optional. We have clarified this in the Methods section (p. 7).	7
Reviewer 3	6	Page 10, under safety, the author should define the difference between grade 3-4 treatment-emergent adverse events (TEAE) and serious TEAE	The grade of each treatment-emergent adverse event (TEAE) describes severity according to NCI-CTCAE v4.0, as stated on p. 8. Each TEAE, regardless of grade, was assessed for seriousness, which is defined by meeting any criterion on a pre-defined list. We added the criteria defining a serious TEAE to the Methods section (p. 8).	8
Reviewer 3	7	Page 11, table 2, worst grade AE. The number of drug-related treatment emergent grade 1 and 2 AE are more than treatment-related AE. This does not make sense. The drug-related should be less, not more. The author should	The numbers in Table 2 for grades 1 and 2 TEAEs and drug-related TEAEs are accurate. In this summary table, the number of drug-related events can be larger than the number of treatment-emergent events for a given grade because a patient is counted only once for each category. In the overall summary of TEAEs, a patient is counted once in the category of worst grade of TEAE regardless of relationship to study drug. To find the drug-related events, a subset is first generated for any TEAEs that are drug related, and then the patient is counted once in the worst grade category. For a given patient, the worst grade	13

		review those figures and correct them.	of drug-related TEAEs may be different from the worst grade of overall TEAEs. We added a footnote to the rows for worst grades 1 and 2 AEs in Table 2 to explain.	
Reviewer 4	1	The study reports a single arm clinical evaluation of Regorafenib for the treatment of refractory metastatic colorectal cancer conducted in 11 Turkish cancer hospitals. Previous randomised evidence for Regorafenib (references 5&6) concluded Regorafenib has a modest but statistically significant improvement on overall survival, tempered by an increased number of grade 3 or 4 toxic adverse events. The study was intended as a descriptive, uncontrolled summary of Regorafenib in the Turkish population, primarily to quantify adverse events but also progression free survival.	This is an accurate summary.	
Reviewer 4	2	The outcomes are consistent with those reported on the previous randomised Regorafenib. The report appears an accurate and	Regarding the exception, we have indicated a response below.	

		balanced description of the study and its outcomes are consistent with those planned in its clinicaltrials.gov record (NCT01853319). As an uncontrolled trial the most relevant reporting standards are a mixture of CONSORT and STROBE, but these appear have been met with one exception discussed below.		
Reviewer 4	3	The comments below are minor, but my main question concerns the importance of this to the journal readership. The focus of this is of limited interest to a international, general medical journal: it is hard to know what this adds. My own background (medical statistician) means I am unable to answer this definitively but I believe the new information presented in this trial is of very limited interest. At the least, some justification for why the two international trial findings (plus other referenced studies) do not address	The randomized, international CORRECT and CONCUR trials assessed the safety and efficacy of regorafenib for patients with treatment-refractory mCRC. Although the CORRECT trial had study sites in Turkey, no patients from Turkey were randomized in CORRECT. In addition, CONCUR was carried out in Asia and did not include Turkey, and Turkey was not included in the large, international, single-arm CONSIGN study. Therefore, REGARD was designed to assess the safety of regorafenib and estimate PFS in a Turkish population. In addition, the study provided access to regorafenib to 100 patients in Turkey prior to market authorization. We have highlighted and clarified the importance of the study in the Introduction (p. 6) and Discussion (p. 17, 19) sections.	6, 17, 19

		safety & efficacy for the Turkish population - or in other words, explanation on why the Turkish population is in some sense different - would help.		
Reviewer 4	4	No sample size justification is presented. The number of patients (100) does seem a reasonable choice for this design and could be retrospectively justificed on grounds of precision (standard error) but the manuscript does not meet generic reporting guidance.	We agree with the reviewer that 100 patients was a reasonable choice based on the size of the country. Based on a similar ongoing global study (CONSIGN) in which Turkey was not a participant, study drug was allocated for 100 patients in Turkey. No statistical assumptions were made. We added a statement to the Methods section to clarify (p 9).	9
Reviewer 4	5	The two previous randomised trials assessed progression using RECIST criteria whereas this has used a more loose definition. Again this may be reasonable since progression free survival is a secondary interest, but should be noted more clearly as a limitation.	Text was added to the Discussion section (p. 18) to highlight the limitation that the response criteria, like the timing of tumor evaluations, depended on each institution's best standard of care.	18

VERSION 2 – REVIEW

REVIEWER	Chaoyuan Kuang University of Pittsburgh, United States
REVIEW RETURNED	27-Dec-2019

GENERAL COMMENTS	Manuscript is appropriate for this venue, prior concerns were addressed and limitations of study design/interpretation are acknowledged in the discussion.

REVIEWER	Mike Bradburn University of Sheffield
REVIEW RETURNED	09-Dec-2019

GENERAL COMMENTS	The authors have addressed most of the previous comments. I remain unclear what information this adds - specifically, why the Turkish population is of particular interest for its own study - and it is unfortunate that no justification was given for the sample size. Nevertheless the report appears a thorough and balanced description of the study findings.

VERSION 2 – AUTHOR RESPONSE

Reviewer 4	1	The authors have addressed most of the previous comments. I remain unclear what information this adds - specifically, why the Turkish population is of particular interest for its own study - and it is unfortunate that no justification was given for the sample size. Nevertheless, the report appears a thorough and balanced description of the study findings.	Although regorafenib was evaluated in three international studies (CORRECT, CONCUR, and CONSIGN), no patients from Turkey were randomized or treated in those studies, as currently stated on pages 6 and 17. As a phase IIIb trial, REGARD was designed to provide additional information on the safety and efficacy of regorafenib in patients with treatment-refractory mCRC (in addition to the data that were generated for regulatory approval) to help further characterize how regorafenib should be used. While the results of REGARD support the findings of the international studies, they were generated in a new context and patient population, and therefore provide additional information	9

			about the use of regorafenib in mCRC. REGARD also enabled patients in Turkey with mCRC and disease progression on all available standard therapies the opportunity to receive an additional line of treatment before regorafenib market authorization. These points are currently stated throughout the manuscript, on pages 3 (abstract, objectives), 6 (introduction), and 16, 17, and 19 (discussion). Determination of the sample size was based on the demand for the study drug and the available supply. We have revised the statement in the manuscript to clarify this (p. 9).	
1	1	Manuscript is appropriate for this venue, prior concerns were addressed and limitations of study design/interpretation are acknowledged in the discussion.	We thank the reviewer for this comment.	