Author's response to reviews

Title: Population-based SEER trend analysis of overall and cancer-specific survival in 5,138 patients with gastrointestinal stromal tumor

Authors:

Ulrich Güller (Ulrich.Gueller@kssg.ch)
Ignazio Tarantino (Ignazio.Tarantino@med.uni-heidelberg.de)
Thomas Cerny (Thomas.Cerny@kssg.ch)
Bruno M Schmied (Bruno.Schmied@kssg.ch)
Rene Warschkow (rene.warschkow@kssg.ch)

Version: 3
Date: 28 May 2015

Author's response to reviews: see over
Dear Dr. Holcomb

We hereby submit the revised version of the manuscript "Population-based SEER trend analysis of overall and cancer-specific survival in 5,138 patients with gastrointestinal stromal tumor" to be considered for publication as an original article in Your Journal.

We found the reviewers` comments very helpful and believe that these comments and our revisions have further improved the quality of this manuscript. All changes in the amended manuscript are bolded, italicised, and underlined.

All authors of this research paper have directly participated in the planning, execution, or analysis of this study. The revised manuscript has been seen and approved by all authors.

Thank you for your time and attention to our revised manuscript.

Sincerely Yours,

Dr. med. René Warschkow, MD, MS
Attending Surgeon

Prof. Dr. med. Ulrich Güller, MD, MHS
Professor of Surgery

On behalf of all authors
Reviewer 1:
Reviewer’s report
Title: Population-based SEER trend analysis of overall and cancer-specific survival in 5,138 patients with gastrointestinal stromal tumor
Version: 2 Date: 26 March 2015
Reviewer: Jordi Rubió-Casadevall
Reviewer’s report:
Major compulsory revisions
This study is an analysis of survival in GIST patients using SEER database in a time period that included years in which the definition of GIST was not clear and before the widespread use of imatinib as treatment of those patients.
I would like to think over some conceptual aspects:
1. In epidemiological studies I think is better to use “relative survival” instead of “cancer specific survival”, as it is difficult to assess the real cause of death without consulting the clinical report of each patient. Relative survival is observed survival adjusted for another causes of death in the population covered, please specify if it has been this the concept used.

Reply to the reviewer: We fully agree with the reviewer that relative survival is an interesting concept. However, the SEER database provides both overall survival and cancer-specific survival as endpoints. The cause of death for cancer specific is unequivocally coded in the SEER database according to the clinical report for each patient and used for this project. We therefore refrained from using the relative survival approach. We fully agree with the reviewer that this issue must be clarified in the manuscript and described in more detail. We have addressed this issue in the methods section on page 5 of the amended manuscript.
2. I believe that the increase in survival in non-metastatic GIST is not real and the reason for the trend in survival in the years before imatinib is known (line226). Before the consensus of 2001 (Fletcher et al. Human Pathol 2002; 33: 459-465) some tumors that are today considered low or very low GIST were diagnosed by pathologist as “leiomioma” or other related terms, not malignant. As considered not malignant, were registered in SEER and another Population-based Cancer Registries. When all tumors suspected to be GIST had been classified performing CD117 immunostaining, almost 25% had been missed by Registries (Rubió et al. Eur J Cancer 2007; 43: 144-148). The inclusion in statistics of these tumors after 2001 has falsely increased incidence and survival (Rubió et al. Clin Trasl Oncol 2014; 16: 660-667). This may be taken into account by authors and somehow included in discussion.

Reply to the reviewer: We fully agree with the reviewer that the increase in survival of non-metastatic GIST patients is confounded by including other pathologies than GIST. We have addressed this issue in the discussion section on page 10.

3. The misclassification of leiomyosarcomas as GIST (line 227) it is not an important bias as the 80% of sarcomas of gastrointestinal tract are GIST, but could be important to discuss that in the first epidemiological study of GIST incidence performed in SEER (Tran et al. Am J Gastroenterol 2005; 100(1): 162-168), CD117 immunostaining was not mandatory to classify this type of neoplasm and incidence was consequently lower than other studies. With a large discussion about the limits and bias of this study, I think is a good epidemiological analysis and suitable to publish.

Reply to the reviewer: We thank the reviewer for this thoughtful comment and have addressed this issue on page 10 in the discussion section of the amended manuscript.

Level of interest: An article of importance in its field
Quality of written English: Acceptable
Statistical review: Yes, but I do not feel adequately qualified to assess the statistics.
Declaration of competing interests: I don’t have any competing interest
Reviewer 2

Reviewer’s report

Title: Population-based SEER trend analysis of overall and cancer-specific survival in 5,138 patients with gastrointestinal stromal tumor

Version: 2 Date: 31 March 2015

Reviewer: Jens Eickhoff

Reviewer’s report:

The authors of this manuscript conducted a trend analysis assessing survival outcomes in a large cohort of GIST patients over a 11-year time period using the SEER database. The analysis provides compelling evidence of a statistically significant and clinical relevant increase in overall and cancer-specific survival from 1998 to 2008. Overall, it is a well written manuscript addressing a highly relevant topic. There are no major methodological flaws.

Major Compulsory Revisions:
- Page 7, The authors indicate that “the statistically optimal cut-off value of mitotic count was 5 in 50 high power fields (HPF). For GIST size, the statistically optimal cut-off is 8cm”. How was the optimal cut-off defined/determined? Presumably, this was based on the results of the ROC curve analysis, but there are different ways to define an optimal cut-off. Please clearly describe how the optimal cut-off was defined/determined.

Reply to the reviewer: We fully agree with the reviewer that this issue needs further explanation. Indeed, the statistically optimal cut-off was estimated as part of the ROC curve analysis by maximizing the Youden index (computed as sensitivity + Specificity-1). We have addressed this issue in the methods section on page 5 and in the results section on page 7.

Minor Essential Revisions:
- Page 17, Table 2: The table summarizes the results of both the univariable and multivariable analysis. Table 2 is labeled as “Multivariable survival analysis”. Please change this to “Univariable and multivariable survival analysis”. Please include the covariates which were used in the multivariable analysis in the footnote of the table.

Reply to the reviewer: We thank the reviewer for this comment and have made the according changes to the manuscript in Table 2.

Level of interest: An article of importance in its field

Quality of written English: Acceptable

Statistical review: Yes, and I have assessed the statistics in my report.

Declaration of competing interests: I declare that I have no competing interests.
Reviewer 3:

Reviewer's report
Title: Population-based SEER trend analysis of overall and cancer-specific survival in 5,138 patients with gastrointestinal stromal tumor
Version: 2
Date: 6 April 2015
Reviewer: Adam Olszewski

Reviewer's report:
Guller et al. describe a survival analysis of patients diagnosed with GIST between 1998 and 2011 using the US SEER registry database. Their primary question was well defined as the impact of introduction of imatinib and further TKI's on survival in GIST as measured by population statistics. They conclude a significant improvement in overall and cancer-specific survival in the study period, which they ascribe to the effect of new systemic therapies.

BMC Review questions:
1. Is the question posed by the authors well defined? Yes.
2. Are the methods appropriate and well described? Generally, yes, with some caveats discussed further.
3. Are the data sound? Generally, yes, with some caveats discussed further.
4. Do the figures appear to be genuine, i.e. without evidence of manipulation? Yes.
5. Does the manuscript adhere to the relevant standards for reporting and data deposition? Yes.
6. Are the discussion and conclusions well balanced and adequately supported by the data? Yes, with some reservations.
7. Are limitations of the work clearly stated? Yes, with some reservations.
8. Do the authors clearly acknowledge any work upon which they are building, both published and unpublished? Generally, yes, with some caveats discussed further.
9. Do the title and abstract accurately convey what has been found? Yes
10. Is the writing acceptable? For the most part, although the text needs further editing.

Discretionary Revisions:
1. I would recommend harmonizing the nomenclature of all genes (KIT, PDGFR, BCR, ABL) according to HUGO Nomenclature. Correct “brc-abl” in particular (it should be bcr). Generally, human GENES should be capitalized and italicized, while PROTEINS should be capitalized.
   Reply to the reviewer: We thank the reviewer and have made the according changes in the amended manuscript.

2. Names of generic drugs (imatinib) need not be capitalized.
   Reply to the reviewer: We thank the reviewer and have made the according changes in the amended manuscript.
3. Some language errors make certain sentences incomprehensible. For example: “For selected pairs, size respectively mitotic rate serving as the cut-off value are shown”. I would also not refer repeatedly to metastatic or adjuvant “situation”.

Reply to the reviewer: We thank the reviewer and deleted the sentence about cut-off values. We changed metastatic and adjuvant “situation” into metastatic and non-metastatic GIST throughout the amended manuscript.

4. Results (Page 6) and Table 1. Consider using either cm or mm for the size consistently to avoid confusion. Listing size statistics to accuracy of 0.1 mm is unlikely to be of any value.

Reply to the reviewer: We thank the reviewer and have made the according changes in the amended manuscript in the amended Table 1 and in the results section on page 7.

5. Table 1 lists N=10 patients with “unknown primary surgery”, but then N=25 patients with “unknown surgery” and M1 disease. Please clarify this discrepancy.

Reply to the reviewer: We thank the reviewer for pointing this out. “Unknown primary surgery” was related to the primary tumour and “unknown surgery” was related to surgery of metastases. We made the according changes in the amended Table 1.

6. Figures and Results (Page 6, Line 167, 169): Please name the panels on Figures 2, 3, 4, 5 and refer to specific panels in the text (rather than “Figure 2”).

Reply to the reviewer: We thank the reviewer and have made the according changes in the amended manuscript in the results section and on the figures 2, 3, 4, and 5.

Minor Essential Revisions:
1. Please explain in the Methods (Page 4, Line 108) the actual primary site and histology codes used. This is important because of known issues with identification of GIST cases in the US registries, which the authors did not discuss as a limitation of the study. As reviewed e.g. in http://seer.cancer.gov/seering/index.php?page=view&id=20100014&type=q only “malignant GIST tumors” are reportable to the SEER registry, and it is unclear whether the registry captures all GIST cases appropriately, and how this had worked out historically, before the specific ICD-O-3 code for malignant GIST was introduced in 2001. Perez et al. (J Am Coll Surg. 2006 Apr;202(4):623-9) showed a 25x increase in reported GIST incidence from 1992 to 2002 in the SEER data, which is almost certainly related to reclassification of various tumors as GIST. This may affect survival rates and make survival trends in this analysis unreliable.

Reply to the reviewer: We thank the reviewer for raising this point. We added the used histology codes (8935 and 8936) in the methods section on page 4. We also added a short paragraph about the limitation of the present investigation concerning the identification and misclassification of GIST in the SEER database and added references to the mentioned study by Perez et al. on page 10.
2. The Methods state that “Chi-square statistics and t-tests were used to compare proportions and continuous variables”, but I see no analyses that would be using those tests in the paper. Moreover, the authors state that “Continuous data are expressed as median and interquartile range”, but then (Page 6, Line 152) describe age as mean and SD. Please explain or correct. 

We thank the reviewer for pointing this out. The reviewer is right and hence we deleted the sentence about these tests and are now only using the median and interquartile range throughout the manuscript (see results on page 6).

3. This analysis shows no difference in prognosis between gastric and small intestinal GIST, which is different than in prior studies. For example, the Gold criteria show significantly worse prognosis for tumors of the small intestine. Please discuss this discrepancy.

Reply to the reviewer: We thank the reviewer. It is not entirely clear why small bowel GIST do have similar prognosis as gastric GIST in the present investigation. We have briefly addressed this issue in the discussion section on page 12.

4. Figure 2: all graphs are cut off at 5 years, although the 5-year survival rates are the main endpoints reported in the results. Please extend the KM graphs to the entire available follow up, or at least to 7 years so that the reader can assess stability of KM curves around the 5-year time point.

Reply to the reviewer: We thank the reviewer for this hint and expanded the Kaplan Meier curves in Figure 2 to seven years.

5. Figure 2: Please harmonize the Y-axis for clarity – the axis in the two top panels are cut off at 50%, which is confusing and exaggerates the survival differences. The same pertains to top panels in Figure 5.

Reply to the reviewer: We thank the reviewer for this comment. The according changes have been made in Figure 2 in the amended manuscript.

6. Page 7, Line 182. The authors state that “For GIST size, the statistically optimal cut-off is 8 cm”. Please explain this statement. What was the purpose of this analysis and for what purpose is this result optimal? If this is “optimal”, why did you decide NOT to use it in your own analysis (you use the 5 cm and 10 cm cutoffs instead). The ROC curves show the trade-off of sensitivity and specificity and the “optimal” assignment depends on many factors, e.g. number of categories desired, whether the high-risk or low-risk group is more important to identify etc.

Reply to the reviewer: The reviewer is raising an important issue. As already requested by Reviewer 2, we have now clarified the estimation of the cut-off value in the methods section on page 5. This estimation was done by maximizing the Youden index. The rationale to do so is to find a discrete value which optimally dichotomizes a cohort. We are fully aware about the methodical drawback about this procedure which is indeed a data-driven estimation. However, we are not aware regarding a better method for cut-off estimation for survival data. The reason not to use the “statistically optimal” cutoff was to keep the results comparable with the usually applied 5 cm and 10 cm cutoff values which are implemented in guidelines. However, we discuss this issue in the amended manuscript in the discussion section on page 12.
Major Compulsory Revisions:
1. Methods (Page 5, Line 141). Please explain how the survival rate was “extrapolated” by Cox regression. Cox regression is a semi-parametric model for (log-)hazard ratio, not a survival probability at some time point, so this kind of modeling is unusual. Additionally, please clarify that you did not include your extrapolated rates in the trend analysis, because this would be fallacious.

Reply to the reviewer: We thank the reviewer for this valuable hint. First of all, we did not include the extrapolated rates in the trend analysis. For the extrapolation, the hazard for an individual is estimated with a certain covariate vector $x$ based on the baseline hazard from the Nelson-Aalen estimator. To estimate the $\beta$, the partial likelihood is maximized. The $\beta$ from this part of the analysis is then used with a covariate vector for which the extrapolation has to be done. We added a short sentence about extrapolation in the methods section on page 6.

2. Methods: (page 5, line 128): please explain for the reader how size (radiology? gross pathology) and mitotic count is recorded in the SEER data. Is mitotic count a continuous factor like size, or is it categorized?

Reply to the reviewer: The reviewer is raising a valid point. Both size and mitotic count are coded as continuous variables with some special values reserved for statements e.g. “less than or equal to 5 mitoses per 50 HPF” which is coded as “990”. Non-numeric values were treated as missing. According to the SEER coding rules, size is recorded as the size from pathology and if not available from radiological estimation. We added a notice in the methods section on page 4.

Data from Table 1 look like there were no missing values for mitotic count. Is that truly the case? Most “site-specific factors” in cancer registry data are known to have a significant proportion of missing values. Consider generating panels on Figure 3 for mitotic count like you did for tumor size so that the reader can understand the distribution of the variable and the relationship with survival.

Reply to the reviewer: We thank the reviewer for this comment. The mitotic count was systematically recorded only after 2009. Hence, as correctly pointed out by the reviewer, this information is lacking for the majority of patients. We clarified this issue in Table 1 in the amended manuscript. The number of events was too small to reproduce the upper panel for the mitotic index.
3. Methods: The endpoints of the study need better description and definition. The title advertises analysis of OS and CSS and both are discussed in the abstract. The Methods do not define any endpoint. How was cancer-specific survival determined? The entire section “Multivariable analysis” makes no comment about CSS analysis at all, although it refers to the table that presents the results.

Reply to the reviewer: We thank the reviewer for this hint and added a definition of the survival endpoints in the statistics section on page 4 and 5. In the results section on page 8 we are now also commenting on the cancer-specific survival.

It would be sound to have a symmetric analysis of both endpoints (OS and CSS) also with regard to the ROC tumor size/mitotic count analysis and the trend—are they prognostic for CSS too? At similar or different cutoffs? At minimum the authors should discuss their conclusions from OS and CSS analysis. The entire trend analysis

Reply to the reviewer: We thank the reviewer for this point. We have added a Figure for ROC analysis of the overall survival and another figure for the trend analysis of the overall survival to present a symmetric analysis of both endpoints. We are now referring to both endpoints (OS and CSS) in the results section. Please see also Figure 5 and Figure 7 in the amended manuscript.

4. The section on pediatric GIST in my opinion should be removed. Discussing mortality of a single patient allows for individual patient identification which is against SEER Data Use Agreement, which states “I will not present or publish data in which an individual patient can be identified. I will not publish any information on an individual patient, including any information generated on an individual case by the case listing session of SEER*Stat. In addition, I will avoid publication of statistics for very small groups”. Describing statistics like mean age and SD in a group specifically defined by age range cutoffs is anyway of dubious value.

Reply to the reviewer: Thank you very much for this hint about anonymity. We fully agree with the reviewer and the section on pediatric GIST has been removed from the amended manuscript.
5. The trend analysis, the primary objective of the study, needs a more consistent presentation and discussion. The title of Figure 5 specifies CSS, while the manuscript text states that the figure represents OS (Page 7, Line 195). Furthermore, hazard ratios from Table 2 and graphs from Figure 5 indicate that any survival improvement occurred only between the 1998-2002 and 2003+ cohort. There is no evident significant improvement between the three subcohorts 2003/5, 2006/8 and 2009/11. This unfortunately is not easy to see because the authors chose to employ a very general likelihood-ratio and Spearman’s correlation tests, which produce P-values, but no actual measures of the trend. In a very large dataset like SEER, P-values are commonly spuriously “significant”, and a more informative epidemiologic measure would be better in order (such as, for example, annual percentage change).

Reply to the reviewer: Thank you very much for this valuable hint. Also considering your comment on “symmetric” analysis of overall and cancer-specific survival, in the amended manuscript there is now one figure for the overall and a second figure for the cancer-specific survival. In the results section on page 9 both figures are now correctly referenced. Also, we estimated the annual percentage change for overall and cancer-specific three-year survival and added it in the results section on page 8 and 9.

The authors correctly discuss that the survival improvement occurred in the period of 1998-2001, which is before imatinib use, but considering no evident improvement afterwards, the entire speculation about the effect of imatinib is dubious. Although the effect of imatinib on survival in clinical trials is without doubt, concluding from these data (which have no indication of any imatinib use) that it was a “cornerstone in the history of cancer care” is arguably a gross exaggeration. The entire effect on survival could be due to reclassification of GIST cases according to new molecular diagnostic criteria developed in 2001.

Reply to the reviewer: We thank the reviewer for these valuables comments. We fully agree with the reviewer that reclassification of GIST impacted the prognosis of our patient cohort. We have expanded on this important issue in the discussion section on page 10. Also, we agree that SEER does not provide data on imatinib use and that advantages imatinib use in clinical trials may not translate to the same extent in population based patient cohort. Hence, we have worded this section more carefully.

6. Discussion, Page 9 Line 241: please correct the erroneous statement that regorafenib “resulted in an overall survival benefit” in “third line treatment”. The referenced GRID study showed no significant difference in overall survival (admittedly with some twists in the recent updated analysis).

Reply to the reviewer: We thank the reviewer for pointing this out. We have changed this in the amended version of the manuscript.
7. Additional limitations of the study warrant discussion:
   a. The authors exclude all cases of GIST which were not patient's first malignancy, likely
      in order to use the “cause-specific survival” as defined in the SEER (although they do
      not motivate this exclusion at all in the paper). However, this exclusion eliminates 1,067
      cases, almost 17%, and may skew the study population by preferentially excluding older
      patients (who have a higher cumulative chance of having an incidental other screening-
      detected cancer). This might result in a systematic bias in the survival estimates and
      trends, especially if the exclusion correlates with calendar years.

      Reply to the reviewer: We thank the reviewer for this interesting remark. As al-
      ready correctly assumed by the reviewer, we excluded patients with other malig-
      nancies preceding the GIST in order to use the cancer-specific survival. The ra-
      tionale for the exclusion of these patients is now explained in the methods section
      on page 4. Additionally, we performed a sensitivity analysis with these 1,067 pa-
      tients included. We attached this analysis to the revision (file Secondary-
      Malig.pdf). As the reviewer can see, the inclusion of these patients did not signifi-
      cantly change the results. The only significant difference was a low predictive
      value of tumor size for overall survival only. Also, we did not find a significant
      correlation between the exclusion rate and the calendar year despite a tendency
      (please see the file SecondaryMalig.pdf).

   b. The authors discuss trends in survival by stage of disease (metastatic vs. non-
      metastatic), but provide no analysis of stage migration. However, PET scans have be-
      come a standard tool in evaluation of GIST in the early/mid-2000’s, potentially leading to
      stage migration which would bias trend analysis because of the Will Rogers phenome-
      non.

      Reply to the reviewer: We thank the reviewer for pointing this valuable comment.
      We have added a short paragraph on the change of imaging over time and the po-
      tential for the Will Rogers phenomenon in the discussion section on page 10.

   c. The authors should at least acknowledge the study by Woodall et al. (JAMA
      Surgery 2009), which analyzed GIST tumors using the same data source and
      had several similar conclusions: size cutoff for poor prognosis of 7 cm.

      Reply to the reviewer: We fully agree with the reviewer and have acknowledged
      the publication by Woodall et al. in the amended manuscript on page 11.

Level of interest:
An article whose findings are important to those with closely
related research interests

Quality of written English: Needs some language corrections before being published

Statistical review: Yes, and I have assessed the statistics in my report.

Declaration of competing interests:
I declare that I have no competing interests