Author’s response to reviews

Title: Resectable adenocarcinomas in the pancreatic head: The retroperitoneal resection margin is an independent prognostic factor

Authors:

Arne Westgaard (arne.westgaard@medisin.uio.no)
Svetlana Tafjord (svetlana.tafjord@rikshospitalet.no)
Inger N Farstad (inger.nina.farstad@rikshospitalet.no)
Milada Cvanarova (milada.smaastuen@rikshospitalet.no)
Tor J Eide (tor.eide@rikshospitalet.no)
Oystein Mathisen (oystein.mathisen@rikshospitalet.no)
Ole PF Clausen (ole.petter.clausen@rikshospitalet.no)
Ivar P Gladhaug (ivar.gladhaug@rikshospitalet.no)

Version: 7 Date: 17 October 2007

Author’s response to reviews: see over
Dear Dr. Edmunds,

Thank you for the thorough review on the statistics delivered by reviewer 3. Our manuscript has now been adjusted and improved based on his suggestions. We believe that we now comply with all the comments made by the reviewers, and hope that the manuscript will be found suitable for publication in your journal.

We would suggest that results from testing of the proportionality assumption, interaction tests, and unadjusted survival analysis, as required by reviewer 3 (referred to as additional file 1, 2, and 3, respectively), might be provided as additional files, since they were only used to verify our statistical methods.

Below is a point-by-point reply including a description of the changes made in our manuscript:

**MAJOR COMPULSORY REVISIONS**

Comment 1: *Statistical analysis section, second paragraph states that “the group with the best prognosis” was used as the reference group. This does not appear to accurately describe the methods that were used, and is probably not the best approach. A better approach is to use the largest group as the referent, which appears to be the method that was actually used for origin of tumor.*

Response: As detailed in additional file 3 (“Unadjusted analysis of histopathologic prognostic factors”), the ampullary group had in fact the longest median survival. For all covariates, the best prognostic group in unadjusted analysis was chosen as referent in adjusted analysis, as stated in our manuscript. We assumed that choosing the best prognostic group was the most meaningful from a clinical point of view and for interpretation of the output. Importantly, the choice of referent does not affect the level of significance, i.e. the overall p-value, for the variable in question. (See also our response...
to comment 12 below). For tumour origin, the referent was also the largest group, and we agree with reviewer 3 that this was thus a good choice as reference group.

Comment 2: *The statistical analysis section indicates that, “Hazard ratios were proportional for all covariates.” I assume you mean that hazards were proportional. You need to indicate (in the methods section) how proportionality was evaluated and present the results of the evaluation in the results section.*

Response: We have now corrected “hazard rations were proportional” to “hazards were proportional”, as correctly pointed out by reviewer 3. The proportionality assumption necessary for fitting the Cox models was evaluated by testing goodness-of-fit based on martingale residual processes, using R version 2.3.1 (open source statistical software, http://www.r-project.org). In addition, we have now included in additional file 1 results from comparison of hazard ratios in the overall base model 1 with hazard ratios obtained by stratification by each covariate. Stratification did not reveal any significant changes in hazard ratios, thus we concluded that hazards were proportional. Graphically, this may be inspected by the included log minus log plots. For simplicity, perioperative mortality (four cases) was excluded in these plots (but not in the tabulated results of proportionality testing).

Comment 3: *The methods indicate that the models were constructed using a stepwise approach. If that method was used, it is not clear why base model 1 includes covariates that are not statistically significant.*

Response: All covariates included in the base models were significantly associated with survival in unadjusted analysis, and thus included in adjusted analysis, as detailed in additional file 3: Unadjusted analysis of histopathologic prognostic factors.

Comment 4: *It is probably not appropriate to exclude the seven patients with multiple margin tumor involvement.*

Response: We excluded the seven cases that had multiple margin involvement in order to evaluate the importance of the individual resection margins, since these cases caused violation of the statistical independence assumption. The seven patients with multiple margin involvement had advanced cancers (T3 or T4 stage) with frequent lymph node metastasis (5 of 7), large tumour diameter (mean 3.1 cm), vessel involvement (6 of 7), perineural infiltration (7 of 7), and areas with poorly differentiated tumour (5 of 7). Thus, these cancers were not typical for our study population. Furthermore, since almost all covariates were unfavourable for these seven patients, assessment of the individual importance of covariates might be obscured by their inclusion. However, adjusted analysis using all 114 patients as suggested by reviewer 3, gave very similar results, and we agree that it is thus unnecessary to exclude the seven patients with multiple margin involvement from the analysis. We have therefore now included all 114 patients in both base models, modelling the resection margins collectively and individually, respectively, as well as added to the text in the statistics section that a separate analysis excluding these 7 cases resulted in the same covariates in the final adjusted model as in the analysis including all 114 patients. (See also our response to comments 7 and 8 below.)
Comment 5: The presentation of the association of resection status by tumour origin (page 9 and figure 5). The discussion of the results of your current analysis (bottom of page 12) seem a bit strong, given that the sample size within each origin site is small, so the difference seen is probably not statistically significant. The appropriate way to evaluate this question is to fit a model with origin and resection status and a model with origin, resection status and the interaction of origin and resection status. A likelihood ratio test comparing those models will evaluate whether there is significant effect modification of resection status by origin.

Response: We agree that the sample size is not sufficient to demonstrate whether there is an interaction between resection status and overall tumour origin when all four groups (ampullary, duodenal, distal bile duct and pancreatic) were included. However, the change in hazard ratio for tumour origin from unadjusted to adjusted analysis indicated an interaction. When tumour origin, resection status and their interaction were modelled simultaneously, the interaction was not statistically significant (p=0.10), as reviewer 3 suspected. This was further confirmed by likelihood ratio test (chi squared=6.07), which was not statistically significant (0.10<p<0.20). However, selecting only the two largest groups for which the sample size was sufficient, the ampullary and the pancreatic groups, an interaction was clearly demonstrated graphically (figure 5) and by Cox analysis using tumour origin (pancreatic versus ampullary), resection status and their interaction (p=0.009). Information about this has now been added to the statistics and results sections of the manuscript, and a table with further details has been included as additional file 3.

Comment 6: Table 1 should be expanded to include a more complete description of the study population, including, at a minimum, the covariates in base model 1.

Response: In our first revision of this manuscript, table 1 included the data as requested also by reviewer 2 (however, reviewer 1 found the additional data superfluous). In our opinion, inclusion of a description of the covariates by tumour origin increases the quality of the manuscript, and previous table 1 has therefore now been reintroduced in the manuscript to comply with the request by reviewer 3.

Comment 7: Base model 2 (table 2) includes distal bile duct margin as a covariate. From Table 1, it appears that this variable defines a group with only 2 patients. It does not make sense to model such a small group.

Response: We agree and have now omitted this covariate from adjusted analysis. The p-values for the retroperitoneal margin (p=0.045) and the pancreatic neck transection margin (p=0.135) in the new base model 2 (presented in table 2B of the revised manuscript) are essentially the same as before excluding the distal bile duct margin from the base model (p=0.051 and p=0.187, respectively, presented in table 2B of the previous version of the manuscript).

Comment 8: My interpretation of the results of base model 2 is quite different than the authors. I don’t think the results argue for evaluating retroperitoneal margins as a separate entity. The hazard ratio for involvement of the pancreatic neck is virtually identical to that for retroperitoneal and distal bile duct also has HR greater than one (although as noted above, it is based on too small a sample to mean anything). I would
conclude that the type of margin is not important, i.e., that microscopic involvement of the margin increases mortality rates. I would not single out retroperitoneal margins.

Response: We agree that microscopic involvement of any margin increases mortality rates. In our manuscript, we have chosen to focus particularly on the retroperitoneal margin due to its clinical importance for both histopathologic evaluation and surgical technique. When performing the surgery, specimen sampling from the pancreatic neck and distal bile duct margins is more easily obtainable than from the retroperitoneal margin. It is of crucial importance that the surgeon at first attempt performs meticulous dissection at the retroperitoneal margin, skeletonizing the vessels, since there will be no second chance of obtaining a margin free resection if tumour is left behind in this location. Proper surgery thus excludes the possibility of obtaining frozen section specimens at this margin. However, the pancreatic neck and distal bile duct margins may be examined by frozen sectioning during surgery and re-resection may be performed as part of the same procedure until free margins at these sites are obtained. In the present study, although involvement of the pancreatic neck transection margin was also statistically significant in unadjusted analysis, and the HR for this covariate was greater than one in adjusted analysis, the p-value was higher than our predefined limit for statistical significance, and the covariate was thus excluded by standardized variable selection. This does however not imply that involvement of this margin is not clinically important, and we agree that a larger sample could demonstrate statistically significant p-values also for this covariate in adjusted analysis. Importantly, we did not wish to conceal the fact that involvement of any resection margin increases mortality rates, and indeed, the results from this analysis was included in the first version of our manuscript (but omitted since reviewer 1 commented that the amount of tabulated data was too extensive). We have now re-introduced the results from modelling the resection margins collectively (table 2C), as indicated by reviewer 3.

MINOR ESSENTIAL REVISIONS

Comment 9: The use of the term multivariate is not appropriate. I would recommend using ‘unadjusted’ and ‘adjusted’ rather than ‘univariate’ and ‘multivariate’.

Response: We agree that the suggested terms are more appropriate and have altered the manuscript text accordingly.

Comment 10: ‘Survival data was’ should be ‘Survival data were’.

Response: We are aware that the plural usage is also common, and have thus altered the text in accordance with the comment by reviewer 3.

Comment 11: “Fischer’s exact test” should be “Fisher’s exact test”.

Response: This has now been corrected.

Comment 12: In Table 2, please explain the p-value (0.056) associated with the reference group for origin site.
Response: SPSS calculates an overall p-value for the importance of each covariate for the group as a whole. This p-value is assigned to the reference group in the statistics output, although it in fact refers to the group as a whole. Some other statistical packages (eg. STATA and R) deliver only p-values for each group versus reference, omitting the overall p-value from the output (but not from comparison with the other covariates). We believe that it is more informative to include the overall p-value as we have done in table 2, in accordance with the SPSS output. In particular, choosing a different referent would change the individual p-values, but the overall p-value would be left unchanged (p=0.056). As mentioned in our response to comment 1, the ampullary group was the best prognostic group in unadjusted survival. However, both the duodenal and distal bile duct groups had HR less than one in adjusted analysis when ampulla was chosen as referent. The individual p-values for duodenum (vs ampulla), distal bile duct (vs ampulla), and pancreatic (vs ampulla) were thus non-significant, although the overall p-value for tumour origin (p=0.056, included in table 2) demonstrates that tumour origin in fact is borderline significantly associated with survival in adjusted analysis (see also our response to comment 1).

Yours sincerely,
Arne Westgaard, M.D.

and

Milada Cvancarova, M.Sc.
(statistician)