Supplementary material to article by S. F. K. Lubeek et al. "Impact of High Age and Comorbidity on Management Decisions and Adherence to Guidelines in Patients with Keratinocyte Skin Cancer'

## Appendix S3

## Details of the regression models for the correlation of different factors related to adherence to guidelines used in this study

First, a descriptive, univariate analysis was performed. Categorical variables were expressed as numbers (and proportions) and compared using a  $\chi^2$  test or Fisher's exact test, when appropriate. Numerical variables were expressed as medians with ranges and compared using a Mann-Whitney test.

Since the distribution of adherence guidelines was expected to be very skew, and since the relation of age and comorbidity with adherence to guidelines is likely to be confounded by various patient and tumour characteristics (Table II), a multivariable logistic regression analysis was performed afterwards.

As no well-established cut-off for a satisfactory level of adherence to guidelines is known from the literature, a cut-off close to the median was selected and 2 categories were formed ("high" vs. "low" adherence to guidelines). This would yield 2 groups of approximately equal size, which maximizes the power of the

study. A priori estimates of study precision would require extensive knowledge about the correlation between all variables to be included in the model. That information is not available. Therefore, we applied the rule of thumb, that for multivariable logistic regression at least 10 observations per variable in the model are required on the least prevalent of the 2 outcomes. Therefore, at least 360 ( $18 \times 10 \times 2$ ) observations were needed in the case adherence guideline has a 50:50 distribution over "high" vs. "low". To allow for an unbalance we decided to aim for inclusion of 400 patients. The distribution with regard to age is not required to be balanced, but a large unbalance would result in a relatively low precision for the effect of age on adherence to guidelines. Since it became clear during the study that the consecutive inclusion in the younger group was much faster than in the old group, after the inclusion of 336 patients (274 aged  $\leq$  80 years; 62 aged  $\geq$  80 years), we continued by only including patients from the older age group. In addition, secondary multivariate logistic regression analyses were performed to study differences in management decisions found in the univariate analyses in more detail.