Author's response to reviews

Title: A methodological approach to identify external factors for indicator-based risk adjustment illustrated by a cataract surgery register

Authors:

Ursula Hahn (U_Hahn@t-online.de)
Irmingard Neuhann (Dr.I.Neuhann@neuhann.de)
Stephanie Schmickler (st.schmickler@augen-zentrum-nordwest.de)
Frank Krummenauer (Frank.Krummenauer@uni-wh.de)

Version: 4 Date: 15 March 2014

Author's response to reviews: see over
Dear Sirs,

first of all, thanks a lot to the reviewers for the constructive and encouraging comments! Below the point by point reply.

Reviewer: Stefan Sauerland

1. Reviewer: The extremely low model fit is discussed as one of the study’s problem. Nevertheless, it would be unrealistic to expect a model fit of R² >80%. Many previous studies that used multivariate methods to analyze postoperative complication rates unfortunately do not report on the proportion of variance explained or R², but it is not uncommon to find Nagelkerke’s R² in the range of 20% to 40%. To some extent, complications after surgery occur by chance, and this holds true especially for less difficult or less risky types of surgery.

Authors: Fully confirmed! We extended the Discussion on this issue; similarly as the point choices for other characteristics in our method proposal the Nagelkerke bound was also rather understood as a proposal for the recent application. Since we modelled a binary endpoint in relation to rather straight cofactors and explaining variables (i.e., we felt modeling a “simple model”), we really expected higher model fit. Regarding your example, we now also declared the Nagelkerke bound as a method parameter to be pre-specified in advance before any evaluation, but in direct context relation.

2. Reviewer: The authors’ decision to consider the stratified analyses more robust than the multivariate regression model is not straightforward. The stratified analyses also cause problems, because external factors are not independent from each other. In table 4, for example, female gender was found to be a significant predictor for visual rehabilitation. As there are no medical reasons to expect females to recover faster than men after cataract surgery, this statistical finding is simply due to the fact that on average females are older than men.

Authors: After reconsideration of the proposal we still feel confident that univariate stratified analysis can be considered more robust than multivariate modeling because of the more robust analysis methods applied for the respective evaluations: note that multivariate modeling always implies some distribution assumptions even in the logistic regression situation, whereas stratified analysis, for example based on the Mantel/Haenszel method, does not. In summary our view on approach robustness was rather derived from an analysis method perspective, but from a result robustness perspective. Nevertheless, we thereby confirm your statement on the result mechanism in the cataract example.

3. Reviewer: On two occasions, the two methodological approaches come to different conclusions: First, severe farsightedness was identified by statistical significance, but this variable failed to reach clinical relevance (8.3% difference). Secondly, presence
of at least one ocular risk factor was clearly significant in the stratified and multivariate analysis, but this variable did not meet the criterion of clinical relevance (7.9% difference). If the authors had defined a 5% rather than a 10% cut-off value for the criterion of clinical relevance, the results for the two criteria would have produced identical results. Therefore, it might be worthwhile to rethink the 10% threshold in future studies, even if the outcome of interest is a frequent event.

Authors: We agree completely: the cut-off value for the criterion clinical relevance depends on the normative setting and the type of endpoint. We tried to exemplify this relationship in the manuscript by writing: “... if an endpoint such as mortality is quantified by means of an indicator, a notably lower threshold limit value for the deviation in prevalences may be appropriate.” We picked up your consideration by adding a new paragraph in section Discussion / clinical relevance.

4. Reviewer
On page 11, death and endophthalmitis are discussed as outcome variables, because both events have a very low incidence (<1%). Admittedly, the 10% cut-off value would need adjustment, because no subgroup of patients in cataract surgery will ever have a mortality or an endophthalmitis rate of 10% or higher. Rather than proposing a flexible incidence-dependent threshold, however, the authors should deliberate about whether a relative rather than an absolute criterion would be more suitable.

Authors: To be honest, we do not feel, that this relative effect parameterization would lower the context dependency – but it would certainly reduce applicability of our method approach description for clinical users. To our interdisciplinary experience, the excess risk scale usually enables clinicians in a more straight-forward way to propose evidence-based clinically relevant thresholds than does the relative risk scale. Context dependency seems to call for this more explicit scale proposition; a “20% relative risk” threshold sounds less context dependent, but again is usually re-scaled onto the excess risk scale more or less explicitly to relate the effect to a baseline magnitude. Although we see your point, we apologize for retaining the more user-orientated scale proposal.

5. Reviewer: The previous comment shows that the two criteria proposed by the authors are in fact closely related to each other. If the presence or absence of an external factor is associated with a large absolute difference in the outcome of interest, it can be expected that this absolute difference is also found to be significant in stratified or multivariate analysis focused on relative differences. Because of the advantages of multivariate modeling it appears much more valid to rely on the multivariate analysis. Thus, the clinical relevance criterion does not confer any additional information and most probably could be deleted.

Authors: Multivariate modeling is indeed a very comprehensive method. But it has disadvantages: As we pointed out above, we consider the stratified analysis more robust (concerning statistical assumptions for their applicability) in identifying external factors. Furthermore: quality assurance programs in medical care are conducted in
direct interaction with physicians, who directly relate to the concept of clinical relevance due its appealing scale (see above). By its advantages and acceptance in data reporting we decided to retain it as well in the methods choice proposal. In addition, for any sample size calculation a priori to quality assurance studies both criteria – clinical relevance as well as statistical significance – are needed (which should therefore again as well be reflected in the method choice proposal).

6. Reviewer: The dichotomization of external factors appears unjustified. Categorization of continuous data, especially dichotomization, reduces the ability of statistical analysis to explain the outcome of interest. If continuous variables such as age or baseline visual acuity are dichotomized, their influence on outcomes may be underestimated or simply remain undetected. In the context of quality assurance not all data are collected as continuous variables, but age in years is certainly recorded very easily. Moreover, it should be described in the methods section, whether the cut-off value of 80 years was defined prospectively or retrospectively (after data inspection).

Authors: The dichotomization was defined a priori: the database was originally set up as a quality monitoring instrument for the affiliated centers. To meet the user's demands in terms of regular report readability we implemented an easily accessible report systematic for endpoints and for (possible) external factors – which was then requested in terms of consistent binary variable representation, wherever possible, by our clinical partners.

Nevertheless, we fully agree with your statement, that for the purpose of identifying relevant external factors it would have been more appropriate to use continuous variables in the first place, which might have had positive impact on the model fit. We added this aspect to sections Material and Methods / Monika Database and Discussion / statistical significance.

7. Reviewer: On page 11, the term “minute prevalence difference” is used. Apparently, the authors mean a “minor prevalence difference”.

Authors: corrected, thanks!

8. Reviewer: There is a typo in table 5: Please correct “constrasts” into “contrasts”.

Authors: corrected, thanks!
Reviewer: Stephan Kirschner

First of all, thanks a lot for your constructive and encouraging reply!

4. The data used for the article seem to be sound. A further description of a valid data set could contribute to even more distinctness.

Authors: We added further information to describe the data set in section “Material and Methods”.

5. The discussion could be improved discussing a possible dependency of the risk factors as well as an analysis of the centre effect.

Authors: We already addressed this special dependency in the manuscript in the following paragraph in section Discussion/Statistical Significance: “Both approaches must be discussed further in respect of multiple testing. Whereas both approaches accounted for centre heterogeneity by either stratification of adjustment for centres, both methods estimated eight external factor associations with each outcome quality indicator. As a consequence, a total of 2 x 8 associations were tested for significance, formally requiring multiplicity adjustment. However, regarding the exploratory nature of identifying external factors requiring risk adjustment in future analysis/reports, the above results should be considered in terms of locally significant association findings instead of formally correcting their confidence interval levels for multiplicity. Nevertheless, researchers should be aware of this “local” significance interpretation underlying the above significance criterion – and should reduce the number of external factors to be considered for risk adjustment.”

6. The low model fit is discussed. For the scientific community further steps to discover so far unknown risk factors may be useful. The structured analysis of complication data may be a source for further risk factors.

Authors: confirmed – we now mention this promising extension of our methodology in the Discussion section.