Comparing patient groups with respect to emesis frequencies in multi-cycle chemotherapies

Werner Vach
Center for Data Analysis and Model Building, University of Freiburg &
Institute of Medical Biometry and Medical Informatics, University of Freiburg

Andreas du Bois
St. Vincentius Hospital, Dept. of Gynecology, Karlsruhe

Abstract: Adequate analysis of emesis frequencies in studies involving multi-cycle chemotherapies is an open problem in antiemetic research. Difficulties arise from the fact that a patient's outcomes in different cycles are not independent, and we describe the two sources of this dependence. Next we show that the widespread approaches of an analysis cycle by cycle with or without adjustment for previous outcomes are of limited value because both tend to produce trends in group differences due to the dependence of a patient's outcomes. As an alternative we propose a new stochastic model which takes into account the two sources of dependence explicitly and hence allows to describe a group difference by a single parameter. The practicability of this model is shown by the analysis of a data set from an open study.

1 Introduction

In early times of antiemetic research it was widely ignored, that many chemotherapies consist of several cycles. Either only the first (available) cycle of each patient was considered, or cycles instead of patients were used as the basic unit, neglecting that efficacy may change over time and that occurence of emesis in a subsequent cycle might depend on the antiemetic control achieved in the prior courses. In recent years we can find more and more publications addressing the problem of multiple cycles in a systematic manner, looking at cycle-specific vomiting rates and the dependence among the outcomes [3,5,8,9,10,13,14,15]. However, the interpretation of the results is often difficult, as it requires some background knowledge on the used methodology. Especially one must be aware, which results, especially trends, can be expected due to the used methodology.

We try to give some insights into two widespread approaches, the analysis cycle by cycle with and without adjustment for the outcome in the previous cycles. We restrict ourselves to the problem to compare patient groups defined by risk factors, for example the comparison of male and female patients, young and old patients or patients with different chemotherapy regimens. In order to understand the typical patterns appearing in comparing risk groups at different cycles, it is necessary to understand the typical sources of dependence among the outcomes of a single patient. We will try to explain these typical sources and to make them partially visible using data from an open study.

So far the cycle-by-cycle analysis played a prominent role in the literature. However, in analysing risk factors we are also interested in summarising the information from all cycles in one single number, allowing a simple interpretation and allowing to compare the contribution of different risk factors. Especially, using information from all cycles allows a much more powerful assessment of risk factors than using only the first cycle, and hence makes it easier to show the significance of risk factors. For this task we propose a new stochastic model, which allows to describe the difference between two risk groups by an appropriate single value. This is achieved by taking into account the sources of dependence among a single patient's outcomes. The use of this model is illustrated at two examples. It gives also a basis to quantify the typical trends we have to expect in an cycle-by-cycle analysis and to discuss the difficulties in defining the persistence of risk factors.

2 Patients and Methods

We use data from an open multi-center study considering single-agent granisetron in 389 patients receiving 989 chemotherapy cycles. The results of this study are published in part elsewhere [7]. All patients received granisetron 3mg i.v. prior to chemotherapy. An additional dose of granisetron was allowed in the case of vomiting or nausea. Occurrence of vomiting was assessed by a patient's diary, and all patients included in our investigation completed this diary for at least 8 days. We summarise this data to a binary outcome variable for each cycle of a patient, indicating vomiting (yes/no) within 8 days after start of chemotherapy.

Within this population we try to investigate two different risk factors: sex and chemotherapeutic regimen. For the first risk factor we consider male and female patients treated with cisplatin with a dose of at least $50mg/m^2$ on the first day of each cycle. These patients received a variety of additional drugs, partially also on subsequent days. For the second risk factor we consider five groups of female patients, receiving either cyclophosphamid + methotrexate + fluorouracil (CMF), fluorouracil + epirubicin + cyclophosphamid (FEC), epirubicin + cyclophosphamid (EC), carboplatin + cyclophosphamid or cisplatin. Carboplatin was administered in a dose of at least $300mg/m^2$, and the dose of cyclophosphamid was at last $500mg/m^2$. The group of female patients receiving cisplatin is used in both analyses, such that we have overall six patient groups. Patient characteristics of these six groups are summarised in Table 1. We include maximally the first four cycles of each patient in our analysis. Vomiting frequencies for each patient group and each cycle are shown in Table 2. There is a substantial portion of drop outs in each cycle, which is due to different reasons: end of or with-drawl from chemotherapy, change of antiemetic treatment, failure in completing the diary over 8 days and others.

Table 1 and 2 about here

In this paper we will make use of logistic models [1, 4], which have become a widely used tool to analyse the impact of one or several variables (covariates) on a binary outcome variable. In the case of comparing patient groups these covariates are binary variables indicating the membership to a specific group. In comparing two groups, we need one indicator variable x with the values 0 and 1. The essential of a logistic model is to describe the conditional probability that the outcome variable y takes the value 1 given the value of the covariate(s) x, and for this task it uses the so called logit transformation, logit $p = \log \frac{p}{1-p}$, which maps probability values between 0 and 1 to the interval $[-\infty, \infty]$. The simplest logistic model with a binary covariate x reads then

$$\text{logit } \Pr(y = 1|x) = \alpha + \beta x$$

i.e., if $x = 0$, the probability on the logit scale to observe $y = 1$ is α, and otherwise $\alpha + \beta$, such that β is a measure for the difference between patients with $x = 0$ and $x = 1$, i.e. between the two patient groups.

If we have several patient groups described by indicator variables $x_1, ..., x_p$, the simplest logistic model reads

$$\text{logit } \Pr(y = 1|x_1, ..., x_p) = \alpha + \beta_1 x_1 + ... + \beta_p x_p \ .$$

In this paper we will consider logistic models for the binary outcome variable vomiting (yes/no) separately for each cycle. Hence we denote by y_t the outcome of a single patient in cycle t.

In the sequel significance always refers to a level of 5 %, and the reported p-values refer to a two-sided test based on the assumption of a Gaussian law of the standardised parameter estimates.

3 Sources of dependence

It is well known that the occurrence of vomiting in one cycle is highly predictive for the occurrence of vomiting in the next cycle [8,10]. With others words, for one patient the occurrences of vomiting in different cycles are not independent. There exist two different reasons for this dependence and the resulting predictive power. The first reason is that the occurrence of vomiting in one cycle increases directly the risk of vomiting in the next cycle. This is probably mainly due to psychological reasons: the bad experience increases the patient's fear of further bad experience, resulting in an increased risk. The most prominent manifestation of this mechanism is anticipatory emesis. The direct change of risk may be also due to a change in the pathomechanism, i.e. the occurrence of vomiting decreases the biological threshold in the next cycle. The second reason is the existence of an individual predisposition of patients, i.e. some patients have a priori a low risk of emesis, other patients have a high risk. In a specific study such an inter-individual variation may be due to variation of known risk factors like sex, age, chemotherapeutic regimen or prior experience, but even in homogeneous populations there will probably remain an individually varying predisposition due to many unknown risk factors. The manifestation of this individual predisposition in clinical life is, that one can observe on the one hand patients never suffering from vomiting and on the other hand patients suffering from severe emesis in all cycles.

It is possible to make these two sources of dependence visible also in the data of our example, if we look at the frequencies of individual vomiting patterns. Table 3a shows the high predictive value of vomiting in the first cycle for vomiting in the second cycle:

more than two thirds of the patients suffering from vomiting in the first cycle vomit in the second cycle, too, but only a small percentage of the patients protected in the first cycle vomit in the second cycle. Table 3a however does not allow to distinguish between the two sources for this predictive power; the outcome in the first cycle may be predictive because of a direct increase caused by vomiting, or because it gives us information on the individual predisposition, i.e. that we have a low risk patient or a high risk patient. Table 3b considers vomiting frequencies in the third cycle in dependence on the pattern in the first two cycles. Patients suffering from vomiting in both cycles have the highest percentage, patients protected in both cycles the smallest percentage. Patients suffering from vomiting in only one of the two cycles are intermediate. However, patients with vomiting only in the second cycle have a higher risk than patients with vomiting only in the first cycle, which can be only explained by a direct increase of the risk caused by vomiting in the previous cycle. If only the second source of dependence, i.e. an individual predisposition would be present, these two patient groups would have an equal risk of vomiting in the second cycle. Evidence for the existence of an individual predisposition can be concluded from Table 3c. Here we look again at all patients completing the first two cycles and we compare the vomiting frequency in cycle 1 with the vomiting frequency in cycle 2 among all patients protected from vomiting in cycle 1. If only the first source of dependence would be present, we would have to expect similar rates or – if the efficacy of antiemetic treatment decreases – a higher rate in cycle 2. Instead, we observe a decrease. This can be only explained by assuming an individual predisposition, such that 'no vomiting' in the first cycle tells us, that these patients have often a low predisposition and hence a lower risk in the second cycle than all patients together in the first. Later on we will show the significance of the existence of these two sources by fitting an appropriate stochastic model.

Table 3a-3c about here

4 Comparison of risk groups cycle by cycle

The most straightforward analysis of vomiting frequencies arising from multi-cycle chemotherapies is a separate analysis of each cycle. In comparing patient groups, however, the insight achievable by such a strategy is limited, as the difference between the patient groups tend to increase over time. This is due to the first source of dependence described in the last section. If we have a high-risk group and a low-risk group of patients, many patients of the high-risk group suffer from vomiting in the first cycle and hence increase their risk in the second cycle, whereas only few patients from the low risk group suffer from vomiting in the first cycle and hence only few patients in this group increase their risk. Consequently, the difference between the two groups will increase over time. This trend for an increase of the group difference may be counterbalanced by an opposite trend caused by drop outs. Drop outs may be related to a previous outcome, because patients suffering from (severe) emesis may be withdrawn from the study more often than patients protected from emesis. As the previous outcome is predictive, more patients with a relatively high risk in the current cycle are withdrawn than patients with a relatively low risk in the current cycle. This effect decrease the observed vomiting frequency in both patient groups, but to a higher degree in the high risk group. Tonato et al. [17] argued, that in comparison of treatment groups this effect may even lead to a change in the observed superiority of the treatments, i.e. to a change of the sign of the difference between the two groups.

A comparison cycle by cycle can - but need not to - be made by fitting a logistic model separately for each cycle. For cycle t the model reads

$$\text{logit } \Pr(y_t = 1|x) = \alpha_t + \beta_t^M x$$

We denote here the cycle specific regression coefficients with the superscript M, as statisticians call this type of analysis a marginal model [6(Chap. 8)]. The above considerations suggest, that the regression parameters β_t^M increase with t, and that this increase is smaller in the presence of drop outs.

5 Comparison cycle by cycle with adjustment for previous outcomes

As the high predictive value of previous outcomes is well known, several authors [9,10,14] suggested to adjust in an analysis cycle by cycle group comparisons for the previous outcomes. This results in fitting logistic models with the previous outcomes as additional covariates for each cycle, such that the model for cycle t reads

$$\text{logit } \Pr(y_t = 1|x, y_1, ..., y_{t-1}) = \alpha_t + \beta_t^T x + \gamma_{t,1} y_1 + \gamma_{t,2} y_2 + ... + \gamma_{t,t-1} y_1$$

We denote here the regression parameters by β_t^T, because statisticians call such a model a transition model [6(Chap. 10)]. This approach seems to be well suited at the first sight, as it takes into account the major source for the increasing trend of the group difference in the analysis cycle by cycle, i.e. the change of risk due to the previous outcome. Additionally one can show, that drop outs depending on previous outcomes do not introduce any bias. Nevertheless, the approach can be criticised, because by adjusting for the previous outcomes we make it more and more difficult to show a relationship between the risk factor defining the groups and the outcome in each cycle. This is due to the fact, that the previous outcomes are variables playing an intermediate role between the risk factor and the outcome in the current cycle; they are partially caused by the risk factor and simultaneously influence the current outcome. The biometrical literature (e.g. [2, p. 104]) warns against adjusting for such intermediate variables, because it renders more difficult to show the significance of the risk factors. Indeed, in a study by the Italian Group of Antiemetic Research [10] using this approach, most risk factors significant in the first cycle fail to be significant in the subsequent cycles. Hence again the approach is of limited value, because a change of significance is probably only due to the chosen methodology.

A more subtle argument shows that there is not only a decreasing trend of significances, but also a decreasing trend of the group differences measured by the parameter β_t^T. We have mentioned above, that both sources of dependence make previous outcomes predictive. The above model adjust only for a direct change due to the previous outcome, and neglects the existence of an individual predisposition. Hence the resulting estimates of $\gamma_{t,s}$ tend to be too large, because they also include the effect of the omitted individual predisposition. Hence, we have an over-adjustment, and as the previous outcomes are correlated with the risk factors, this leads to a decrease of the estimates of β_t^T over time due to an increasing over-adjustment. Hence we have a decreasing trend in the group differences measured by β_t^T.

6 A new stochastic model

We have seen above that there exist two different sources for the dependence of a patient's outcomes among different cycles. These dependence render the interpretation of an analysis cycle by cycle difficult, because they induce typical trends in group differences and their significance. To avoid such trends caused by the chosen methodology, we propose a new stochastic model, which takes both types of dependence into account explicitly. To describe the individual predisposition, in addition to the observable variables $y_1, ..., y_t$ and x we introduce the unobservable variable l which takes different values for different patients. The direct change of risk due to vomiting in the previous cycle is described by the parameter δ, which is the increase of the probability of vomiting on the logit scale. We additionally assume that the occurrence of vomiting in cycle t has not only an impact on cycle $t+1$, but also on the subsequent cycles $t+2, t+3, ...$ to a smaller extent. Precisely we assume, that vomiting in cycle t increases the risk in cycle $t+j$ by the amount $\delta\gamma^{j-1}$, where γ is an additional parameter with values between 0 and 1. Then the new stochastic model, which describes the probability of vomiting in cycle t given the group membership indicator x, the previous outcomes $y_1, ..., y_{t-1}$ and the individual predisposition l reads

$$\text{logit } \Pr(y_t = 1|x, y_1, ..., y_{t-1}, l) = \alpha_t + \beta_t x + l + \delta y_{t-1} + \delta\gamma y_{t-2} + \delta\gamma^2 y_{t-3} + ... + \delta\gamma^{t-2} y_1$$

where β_t is the group difference in cycle t, and α_t is the probability of vomiting in cycle t for a patient in the first group ($x = 0$) with an individual predisposition equal to 0 and without previous vomiting ($y_1 = 0, ..., y_{t-1} = 0$). As we have taken into account the two sources of dependence it is now – at least in a first approximation – allowed to assume that β_t has a constant value β over time. This is contrary to the analysis cycle by cycle with or without adjustment for previous outcomes, where the parameters β_t^M or β_t^T tend to increase or decrease, respectively, over time. Hence we are now able to describe the difference between two groups in a meaningful and adequate way by a single parameter β. The interpretation of β is, that two patients, one from each group, which have the same individual predisposition, differ at each cycle t with respect to the probability of vomiting by the amount β on the logit scale, given that they have the same vomiting pattern in the previous cycles. Examples for using the model for the task to describe group differences will be given in the next section.

At first sight the new model seems to be of no practical value, as it depends on the unknown value l. However, it is well known in the statistical literature (e.g. Laird & Ware [11]) that nevertheless such a model can be fitted simultaneously to the data of all cycles, if we assume that the individual values l follow a Gaussian law with mean 0 and unknown standard deviation σ. Note that the individual predisposition reflects a sum of effects of many unknown risk factors such that the Gaussian law can be justified by the classical central limit theorem. Details of computing parameter estimates under this assumption by using the maximum likelihood principle are given in Vach et al. [16]. As data in antiemetic research are usually affected by drop outs, it is important to note, that the model can be easily fitted to data with drop outs, too. To assure that the drop outs introduce no bias, we have to assume that any outcome which causes a drop out must be recorded in the data, and it must be excluded that a drop out hides the event which caused it. For example, if a patient suffers from severe emesis in cycle t such that already in this cycle the antiemetic treatment is changed, the value $y_t = 1$ must be still recorded, and only the values $y_{t+1}, y_{t+2}, ...$ of the subsequent cycles are allowed to be missing. So bias due to drop outs can be excluded by appropriate rules in the study protocol.

There are two slight draw backs of the proposed model, which have been detected by the investigations of Vach et al. [16]. First, in estimating the model parameters it may happen, that the maximum likelihood principle produces ambiguous estimates. This ambiguity is related to the fact, that we offer in our model two distinct explanations for the observed correlations among the outcomes, and the data fails sometimes to separate the contributions of the two sources. This implies, that we have to check the uniqueness of the parameter estimates in each specific application. Second, it is possible to improve the efficiency in proving the significance of risk factors by considering a specific function of the parameter estimates. We neglect this further step towards optimality in this paper, the interested reader is referred to Vach et al. [16].

Finally some comments on the intercept parameters α_t are necessary. In order to obtain parsimonious models, one may assume in a naive analysis, that α_t has a constant value α over time after adjustment for the individual predisposition. However, there exists at least one reason to allow a change over time: a possible loss in efficacy of the antiemetic treatment. In order to keep the model parsimonious, we suggest to consider in a standard analysis only a linear change over time, i.e. $\alpha_t = \alpha + (t-1)\Delta$. Hence the model we fit simultaneously to the data of all cycles reads for cycle t

$$\text{logit } \Pr(y_t = 1|x, y_1, ..., y_{t-1}, l) = \alpha+(t-1)\Delta+\beta x+l+\delta y_{t-1}+\delta\gamma y_{t-2}+\delta\gamma^2 y_{t-3}+...+\delta\gamma^{t-2}y_1$$

and we have the six parameters $\alpha, \Delta, \beta, \sigma, \delta$ and γ to be estimated, where σ is the standard deviation of the individual predisposition l.

7 Example

In the first part of the example we analyse the difference between the two sex groups for the cisplatin treated patients. The covariate x has the value 0 for males and the value 1 for females, and the global difference between the two sexes is described by the value of the parameter β. The maximum likelihood estimates together with their standard errors are shown in Table 4. We observe, that globally women have a risk increased by an amount of 1.51 on the logit scale in comparison to men, and this difference is significant ($p = 0.02$). In addition, we can get some insight in the temporal development and the dependence between different cycles. Vomiting in the previous cycle has an impact on vomiting in the current cycle ($\delta = 1.09$), but nearly no impact on subsequent cycles ($\gamma = 0.02$). The variability of the individual predisposition is described by a standard deviation of $\sigma = 1.84$. The magnitude of this variability can be expressed by the expected absolute difference $|l_2 - l_1|$ of the individual predisposition of two randomly selected individuals, which is estimated as $4\sigma/\sqrt{2\pi} = 2.94$, and hence larger than the difference between males and females. However, the estimates of δ, γ und σ show a large standard deviation and hence our knowledge of the direct increase of risk and the variability of the individual predisposition is rather vague. There is no evidence for a change of the intercepts over time ($\Delta = 0.01$), and hence no evidence for a loss of efficacy.

Table 4 about here

In the second part of the example we consider the difference between the five female patient groups. Here we need four binary indicator variables $x_1, ..., x_4$, which indicate membership in the CC, EC, FEC and cisplatin group, respectively, whereas the CMF group serves as a reference group. The model reads now

$$\text{logit } \Pr(y_t = 1 | x_1, ..., x_4, y_1, ..., y_{t-1}, l) = \\ \alpha + (t-1)\Delta + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + l + \delta y_{t-1} + \delta\gamma y_{t-2} + \delta\gamma^2 y_{t-3} + ... + \delta\gamma^{t-2} y_1$$

where $\beta_1, \beta_2, \beta_3, \beta_4$ describe the differences on the logit scale in comparison to CMF. The resulting estimates are shown in Table 5. The hypothesis of no difference between all groups can be rejected with a p-value smaller than 0.0001. The small value of β_1 indicates no great difference between the CMF and the carboplatin + cyclophosphamid group. EC and FEC show both a distinctly increased risk of vomiting in comparison to CMF, but the difference between both is very small ($\beta_3 - \beta_2 = 0.20$). The value of β_4 indicates that the cisplatin patients have the largest risk of vomiting and that this risk is distinctly larger than in all other groups. In interpreting the differences between the chemotherapy groups one has to keep in mind that the compared groups differ also in other characteristics than the regimen. For example, EC is a typical second line therapy, such that the difference between EC and CMF may be due to a difference in the prior experience of these patient groups. Due to the larger sample size we can now show that the values of δ and σ are both significant different from zero ($p = 0.048$ and $p = 0.02$, respectively), and we have now also some evidence for an effect of vomiting in the cycle t on vomiting in the cycle $t + 2$ or later, as γ is away from zero. Again there is no evidence for a change of the intercepts over time.

Table 5 about here

In both examples we tried several start values in the estimation procedure to check for an ambiguity of the estimates, but no hints were found. Note that there was a strict rule in the study protocol that the outcome of any cycle, which started with a treatment with granisetron prior to chemotherapy, has to be recorded in the data, such that in spite of substantial drop out rates we have to expect no bias in our analysis.

8 Quantifying trends in the analysis cycle by cycle

Our considerations on the analysis cycle by cycle above have been on a qualitative level. We were able to argue that typical trends will be observed in this type of analysis, however, we were unable to show, whether such trends are substantial or negligible. With the help of the new stochastic model we are now able to quantify these trends, because the model allows to describe the joint probability law for the outcomes of the patients. Hence we can choose specific values for the six parameters $\alpha, \Delta, \beta, \sigma, \delta$ and γ, which reflect situations of practical relevance, and then we can compute the expected values of the parameters β_t^M or β_t^T, respectively, occurring in the analysis cycle by cycle. Expected value means here the value we have to expect on average in analysing large data sets generated by the law determined by the parameter values.

We start our considerations with the expected values of β_t^M shown in Table 6. We vary in this table the values of δ and σ, describing the two sources of dependence and mainly determining the course of the parameter β_t^M over time, and keep the values of β, γ, α and Δ fixed. We observe the increase described above, and the degree of increase increases with δ, as this implies a high direct increase of the risk due to previous vomiting, and it decreases with σ, as in the case of a highly varying individual predisposition the role of the direct increase is relatively small. Within the considered parameter constellations we observe an increase up to 50 % between cycle 1 and cycle 4. However, this happens only in the case of a small variation in the individual predisposition. The example of the last section suggests that a high variation of the individual predisposition may be typical, and in this case the increase is much smaller. This may explain why so far no results indicating a dramatic increase of group differences have been reported. Note that also in the first cycle the value of β_1^M tends to be smaller than the value of $\beta = 1.0$ in the assumed model. This decrease due to non-adjustment for an individual predisposition is well known in the statistical literature [12,18].

Table 6 about here

We have mentioned above that the increase of group differences over time may be counterbalanced by drop outs, if vomiting patients are withdrawn more often than non-vomiting patients. We are now able to quantify this effect, too. For this we recalculate the expected values of β_t^M for a constellation with drop outs. We assume, that a patient protected from vomiting in a specific cycle has a probability to be withdrawn after this cycle of 10%, and consider different values for the probability to be withdrawn for patients suffering from vomiting. In Table 7 we can observe, that slight to moderate differences in the dropout probabilities have no great impact on the course of the differences over the cycles, and only large differences - here 10% drop out probability for non-vomiting patients and a 50% and higher drop out probability vor vomiting patients - really matter. A change of an increase to a decrease of the course in time can only be observed for the most extreme constellations. This result, which is similar also for other parameter constellations, suggests, that the possible bias in group comparisons due to drop outs has been overestimated so far. Drop outs are of course a source of bias if we do not look at group differences, but at the vomiting frequencies within each group, which we expect to increase in the case of no drop outs, but which may decrease in the presence of drop outs. Drop outs can be also a source of bias in the case of group comparisons, if there is a qualitative difference in the drop out mechanism of the groups, for example in the first group there is no relation between previous vomiting and drop outs, but in the second there is one. However such constellations may be not very probable in antiemetic research.

Table 7 about here

In the next step we consider the trends in the analysis cycle by cycle with adjustment for previous outcomes. Table 8 shows the expected values of the parameters β_t^T for some parameter constellations. We have argued above, that the existence of a varying individual predisposition implies a decrease of the group difference over time. Indeed, if we have no variation of the individual predisposition ($\sigma = 0$), the expected values of β_t^T are constant over time, but otherwise we observe a decrease, which is in the considered constellations maximally about 50% of the value in the first cycle. The value of δ has little impact on the decrease. The figure in Table 8 does not change if additionally drop outs depending on previous outcomes occur, because we condition on the previous outcome.

Table 8 about here

However, the trend in the expected values is only one problem with this approach. The other one is the loss of power due to the adjustment for intermediate variables. To quantify this loss we compute for each cycle the power to show a significant difference, i.e. the probability to reject the null hypothesis $\beta_t^T = 0$ at the five-percent level, if we compare two groups of size 80 and a true difference of $\beta = 1.0$ in our model. We consider in Table 9 the same parameter constellations as in the previous table, and we can observe a loss of power even for those constellations, where the expected values of β_t^T are constant over time. The loss of power is most distinct in the case of a high individual variation combined with a high direct increase of the risk due to previous vomiting. As this must be regarded as typical for data in antiemetic research, we have to conclude, that adjusting for previous outcomes prevents risk factors from reaching significance. This explains at least in part the results of the study of the Italian Group on Antiemetic Research [10].

Table 9 about here

9 The persistence and non-persistence of risk-factors – a matter of definition

It is easy to speak of the persistence of risk factors, however, it is hard to define this term precisely. The characteristic of a risk factor is that it can help us to predict the probability of vomiting of a patient, and hence a first straightforward choice is to define the persistence of a risk factor as the persistence of its predictive value. In the naive analysis cycle by cycle the predictive value is analysed in each cycle, and here the predictive value increases. If we adjust for previous outcomes the predictive value decreases, because we add in each cycle more and more other informative variables. Hence defining the persistence based on the predictive value is not very useful, because non-persistence is the natural case. This does not imply, that such analyses are worthless. They can help the clinician to decide, whether to look only at previous outcomes or also at additional risk factors in the individual prediction of emesis and hence in the individual decision of an optimal antiemetic treatment. In our model we have assumed a priori that the regression coefficient β_t is constant over time, and we believe, that this type of persistence of a risk factor is a justifiable assumption in most settings. Of course, it is straightforward to consider also variants of our model with non-persistence of the risk factor. The simplest one is given by the assumption of a linear change of the difference; i.e. $\beta_t := \beta + (t-1)\Delta_\beta$, such that the model reads

$$\begin{aligned}&\text{logit } \Pr(y_t = 1|x, y_1, ..., y_{t-1}, l) = \\ &\qquad \alpha + (t-1)\Delta + (\beta + (t-1)\Delta_\beta)x + l + \delta y_{t-1} + \delta\gamma y_{t-2} + \delta\gamma^2 y_{t-3} + ... + \delta\gamma^{t-2} y_1 \ .\end{aligned}$$

We have now a single parameter Δ_β measuring non-persistence, but what is its interpretation? We can rewrite the above model in an equivalent manner by considering the values $x = 0$ and $x = 1$ separately. It reads then

$$\text{logit } \Pr(y_t = 1|x = 0, y_1, ..., y_{t-1}, l) = \alpha^{x=0} + (t-1)\Delta_\alpha^{x=0} + l + \delta y_{t-1} + ... + \delta\gamma^{t-2} y_1$$

and

$$\text{logit } \Pr(y_t = 1|x = 1, y_1, ..., y_{t-1}, l) = \alpha^{x=1} + (t-1)\Delta_\alpha^{x=1} + \delta y_{t-1} + ... + \delta\gamma^{t-2} y_1$$

and the two representations are equivalent with $\alpha^{x=0} = \alpha$, $\Delta_\alpha^{x=0} = \Delta$, $\alpha^{x=1} = \alpha + \beta$ and $\Delta_\alpha^{x=1} = \Delta + \Delta_\beta$. Hence we have now different changes of the intercepts in the two risk groups, and $\Delta_\beta := \Delta_\alpha^{x=1} - \Delta_\alpha^{x=0}$ just measures this difference. We have mentioned above that a change of the intercept can be caused by a loss of efficacy, and this is nearly the only explanation. Hence analysing non-persistence of risk factors by adding Δ_β to our model is nothing else than analysing a difference of the loss of efficacy between the risk groups. Although loss of efficacy is a very interesting question, we still hesitate to use the above model for this question. One must be aware, that it is much more difficult to obtain valid estimates for the change of efficacy than for a group difference, because the first requires the analysis of the correlated outcomes of a single patient, whereas the later is based on comparing independent observations. Hence for this step further investigations are necessary, before recommendations can be made.

10 Discussion

In comparing patient groups with respect to vomiting frequencies in multi cycle chemotherapies an analysis cycle by cycle is only of limited value. Without adjustment for previous outcomes group differences tend to increase over time, with adjustment, the power to show a significant difference decrease. Hence the results obtainable by such analyses tend to reflect to a high degree only the chosen methodology. As an alternative we suggest to concentrate in a first step on describing group differences by an appropriate chosen single parameter, which has an interpretation in each cycle. For this task we propose a stochastic model, which takes into account the sources of dependence among a patient's outcomes explicitly, and which hence allows to define such a parameter. The practicability of this approach has been shown by two simple examples comparing sex groups or chemotherapeutic regimens, respectively. We used this model also to quantify possible trends in the analysis cycle by cycle and to discuss the difficulties in defining the persistence of risk factors.

We focused in this paper on the comparison of patient groups defined by risk factors. Comparisons of therapy groups are much more difficult. For risk groups the natural assumption is a constant difference over time, for treatment groups it is much more natural to allow a change of the difference over time, hence this requires a much more careful modeling. As a basic alternative one may focus in a treatment comparison on the complete protection from vomiting over all cycles, or, which is nearly equivalent, on the time until the first cycle with vomiting, which allows to use methods of survival analysis like Kaplan-Meier curves [5,15]. Such methods can be, of course, also used to analyse risk factors, but they totally ignore all information following the first cycle with vomiting, and hence are wasteful of information. As a compromise one may use the number of cycles with vomiting as the outcome measure, but drop outs make it very difficult to analyse this type of data in an adequate manner.

References:

1. Altman DG (1991) Practical Statistics in Medical Research. Chapman & Hall, London, Chap. 12.5
2. Breslow NE, Day NE (1980) Statistical methods in cancer research, Vol. 1 – The analysis of case-control studies. IARC Scientific Publications No. 32, Lyon
3. Chevalier-Evain V, Bonneterre J, Adenis A, Pion JM, Hecquet B (1994) Loss of efficiency of ondansetron-dexamethasone during successive courses in female patients receiving high-dose cisplatin. Bulletin du Cancer 81:219-222
4. Cox DR (1970) Analysis of binary data. Methuen, London
5. de Wit R, Schmitz PIM, Verweij J, de Boer-Dennert M, de Mulder PHM, Planting AST, van der Burg MEL, Stoter G (1996) Analysis of cumulative probabilities shows that the efficacy of $5HT_3$ antagonist prophylaxis is not maintained. Journal of Clinical Oncology 14:644-651
6. Diggle PJ, Liang KY, Zeger SL (1994) Analysis of Longitudinal Data. Oxford Science Publications, Oxford
7. Golaszewski T, Kolbl H, Lahousen M, Staab HJ (1995) Granisetron, ein Antiemetikum für die Behandlung der Zytostatika-induzierten Emesis: Ergebnisse einer praxisorientierten Studie (in German). Wiener Klinische Wochenschrift 107:613-621
8. Italian Group for Antiemetic Research (1993) Difference in persistence of efficacy of two antiemetic regimens on acute emesis during cisplatin chemotherapy. Journal of Clinical Oncology 11:2396-2404
9. Italian Group for Antiemetic Research (1994) Cisplatin-induced delayed emesis: Pattern and prognostic factors during three subsequent cycles. Annals of Oncology 5:585-589

10. Italian Group for Antiemetic Research (1995) Persistence of efficacy of three antiemetic regimens and prognostic factors in patients undergoing moderately emetogenic chemotherapies. Journal of Clinical Oncology 13:2417-2426

11. Laird, NM and Ware JH (1982) Random-effects models for longitudinal data. Biometrics 38:963-974

12. Neuhaus JM, Kalbfleisch JD, Hauck WW (1991) A comparison of cluster-specific and population-averaged approaches for analyzing correlated binary data. International Statistical Review 59:25-35

13. Silva RR, Bascioni R, Giorgi F, Acito L, Giustini L, De Signoribus G, Marcellini M, Menichetti ET, Giuliodori L (1996) Granisetron plus dexamethasone in moderately emetogenic chemotherapy: evaluation of activity during three consecutive courses of chemotherapy. Supportive Care in Cancer 4:287-290

14. Soukop M (1996) Management of cyclophosphamide-induced emesis over repeat courses. Oncology 53:Suppl. 1, 39-45

15. Soukop M, McQuade B, Hunter E, Stewart A, Kaye S, Cassidy J, Kerr D, Khanna S, Smyth J, Coleman R, et al. (1992) Ondansetron compared with metoclopramide in the control of emesis and quality of life during repeated chemotherapy for breast cancer. Oncology 49:295-304

16. Vach W, du Bois A, Schulz K (1997) A stochastic model to analyse risk factors for emesis in multi-cycle chemotherapies. Preprint Nr. 37, Center for Data Analysis and Model Building, University of Freiburg

17. Tonato M, Roila F, del Favero A, Ballatori E (1996) Methodology of trials with antiemetics. Supportive Care in Cancer 4:281-286

18. Zeger SL, Liang KY, Albert PS (1988) Models for longitudinal data: A generalized estimating equation approach. Biometrics 44:1049-1060

Chemo-therapy	Sex	# patients	# cycles	< 45	age 45-60	> 60	diagnoses
CMF	f	167	491	58	74	35	breast cancer (167)
CC	f	40	98	2	16	22	ovarian cancer (40)
EC	f	40	100	10	21	9	breast cancer (40)
FEC	f	20	61	9	6	5	breast cancer (20)
Cis	f	61	128	18	23	20	ovarian cancer (34), endometrial cancer (6), others (21)
Cis	m	61	131	13	31	17	head/neck cancer (24), lung cancer (18), others (19)

Table 1: *Patient characteristics for the six considered risk groups. CMF = cyclophosphamid + methotrexate + fluorouracil, CC = carboplatin + cyclophosphamid, EC = epirubicin + cyclophosphamid, FEC = fluorouracil + epirubicin + cyclophosphamid, Cis = cisplatin*

Chemo-therapy	Sex	Cycle 1		Cycle 2		Cycle 3		Cycle 4	
		n	%	n	%	n	%	n	%
CMF	f	167	17	128	16	110	22	86	21
CC	f	40	20	24	17	20	20	14	14
EC	f	40	35	28	18	20	25	12	42
FEC	f	20	20	17	59	13	69	11	37
Cis	f	61	61	35	66	21	67	11	55
Cis	m	61	38	37	38	21	48	12	42

Table 2: *Vomiting frequencies stratified by patient group and cycle. n = number of available patients, % = percentage of vomiting patients. Abbreviations of chemotherapies are explained in Table 1*

		vomiting 2nd cycle		
vomiting 1st cycle	n	no	yes	%
no	205	173	32	15.6%
yes	64	19	45	70.3%

Table 3a: *Vomiting patterns within the first two cycles for the 269 patients finishing the first two cycles*

vomiting in cycle		n	vomiting 3rd cycle		
			no	yes	%
1	2				
no	no	138	116	22	15.9%
no	yes	24	9	15	62.5%
yes	no	15	9	6	40.0%
yes	yes	28	5	23	82.1%

Table 3b: *Vomiting patterns within the first three cycles for the 205 patients finishing the first three cycles*

		vomiting 1st cycle		
All patients	n	no	yes	%
	269	205	64	23.8%
		vomiting 2nd cycle		
Patients with no vomiting in in the first cycle	n	no	yes	%
	205	173	32	15.6%

Table 3c: *Vomiting frequency in the first two cycles for the 269 patients finishing the first two cycles*

	est	std
β	1.51	0.65
δ	1.09	0.08
γ	0.02	0.58
σ	1.84	0.78
α	-0.82	0.42
Δ	0.01	0.32

Table 4: *Parameter estimates (est) and standard errors (std) from fitting the new stochastic model with a single covariate x representing membership to the sex groups to the data of all cycles for the cisplatin patients*

	est	std
β_1	0.09	0.43
β_2	1.05	0.46
β_3	1.25	0.57
β_4	2.62	0.68
δ	1.10	0.56
γ	0.34	0.35
σ	1.34	0.57
α	-2.10	0.45
Δ	-0.03	0.20

Table 5: *Parameter estimates (est) and standard errors (std) from fitting the new stochastic model to the data of the female patients*

$\sigma = 0.5$				
δ	$t = 1$	$t = 2$	$t = 3$	$t = 4$
0.0	0.95	0.95	0.95	0.95
0.5	0.95	1.03	1.09	1.12
1.0	0.95	1.10	1.21	1.28
1.5	0.95	1.15	1.31	1.43
$\sigma = 1.0$				
δ	$t = 1$	$t = 2$	$t = 3$	$t = 4$
0.0	0.83	0.83	0.83	0.83
0.5	0.83	0.88	0.92	0.94
1.0	0.83	0.93	0.99	1.03
1.5	0.83	0.96	1.05	1.12
$\sigma = 1.5$				
δ	$t = 1$	$t = 2$	$t = 3$	$t = 4$
0.0	0.71	0.71	0.71	0.71
0.5	0.71	0.74	0.76	0.77
1.0	0.71	0.77	0.81	0.83
1.5	0.71	0.79	0.84	0.87

Table 6: *Change of the expected value of β_t^M over time for different parameter constellations in dependence of δ and σ. The remaining parameter values are $\beta = 1.0, \gamma = 0.5, \alpha = -0.5$ and $\Delta = 0.0$.*

p_0	p_1	$t=1$	$t=2$	$t=3$	$t=4$
0.1	0.1	0.83	0.93	0.99	1.03
0.1	0.2	0.83	0.93	0.99	1.03
0.1	0.3	0.83	0.92	0.98	1.02
0.1	0.4	0.83	0.91	0.97	1.00
0.1	0.5	0.83	0.90	0.94	0.96
0.1	0.7	0.83	0.87	0.86	0.82
0.1	0.9	0.83	0.78	0.71	0.64

Table 7: *Change of the expected value of β_t^M over time for different constellations of drop out probabilities for vomiting (p_1) and non-vomiting (p_0) patients. The values of the parameters are $\beta = 1.0, \delta = 1.0, \sigma = 1.0, \alpha = -0.5$ and $\Delta = 0.0$.*

δ	$t = 1$	$t = 2$	$t = 3$	$t = 4$
		$\sigma = 0.0$		
0.0	1.0	1.0	1.0	1.0
1.0	1.0	1.0	1.0	1.0
		$\sigma = 0.5$		
0.0	0.95	0.90	0.85	0.81
1.0	0.95	0.90	0.86	0.83
		$\sigma = 1.0$		
0.0	0.83	0.71	0.62	0.55
1.0	0.83	0.72	0.63	0.57
		$\sigma = 1.5$		
0.0	0.71	0.55	0.45	0.38
1.0	0.71	0.56	0.46	0.40

Table 8: *Change of the expected value of β_t^T over time. The remaining parameter values are $\beta = 1.0, \gamma = 0.5, \alpha = -0.5$ and $\Delta = 0.0$.*

δ	$t=1$	$t=2$	$t=3$	$t=4$
		$\sigma = 0.0$		
0.0	86	87	83	81
1,0	86	84	74	70
		$\sigma = 0.5$		
0.0	83	77	71	66
1.0	83	72	63	54
		$\sigma = 1.0$		
0.0	73	58	44	37
1.0	73	51	42	31
		$\sigma = 1.5$		
0.0	61	38	26	20
1.0	61	32	21	16

Table 9: *Power (in %) to reject the null hypothesis* $\beta_t^T = 0$ *for two patient groups of size* 80 *in dependence of* δ *and* σ. *The remaining parameter values are* $\beta = 1.0, \gamma = 0.5, \alpha = -0.5$ *and* $\Delta = 0.0$. *Results are based on a simulation study with 1000 replications.*