Chapter 8

Performing survival analyses in the presence of competing risks

A clinical example in older breast cancer patients

Nienke A de Glas*
Mandy Kiderlen*
Jan P Vandenbroucke
Anton JM de Craen
Johanneke EA Portielje
Cornelis JH van de Velde
Gerrit-Jan Liefers
Esther Bastiaannet
Saskia le Cessie

**shared first authorship*

Submitted for publication

Abstract

An important consideration in studies that use cause-specific endpoints such as cancer-specific survival or disease recurrence, is that the risk of dying from another cause before experiencing the event of interest is generally much higher in older patients. These so-called competing events are of major importance in the design and analyses of studies with older patients, as a patient who dies from another cause before the event of interest occurs, can obviously not experience this anymore. The aim of this study was to present several clinical examples of research questions in a population-based cohort of older breast cancer patients with a high frequency of competing events, and to discuss the implications of choosing models that deal with competing risks in different ways.

Our results show that in populations with a high frequency of competing events, it is important to consider which method is the most appropriate to estimate cause-specific endpoints. In summary, when calculating absolute cause-specific risks, we demonstrate that the Kaplan Meier method overestimates the risk of the event of interest, and that the cumulative incidence competing risks (CICR) method, which takes into account the competing risks, should be used instead. There are two common used approaches to modelling the association between prognostic factors and cause specific survival: the Cox proportional hazards model and the Fine and Gray model. We discuss both models and show that often in *etiologic* research, the Cox Proportional Hazards model is more appropriate, while in *predictive* research, the Fine & Gray Model is recommended.

In conclusion, in studies with cause-specific endpoints in populations with a high frequency of competing events, researchers should carefully choose the most appropriate statistical method in order to prevent incorrect interpretation of study results.

Background

In order to study treatment efficacy or other exposures in clinical research, large cohorts of patients are often followed during a certain period of time. Frequently, cause-specific endpoints are used in these studies, such as recurrence, cancer-specific mortality or cardiovascular mortality[1]. For these endpoints, statistical methods that assess the time to an event such as the Kaplan Meier method or the Cox Proportional Hazard model are frequently used[2;3].

An important consideration in studies that use these cause-specific endpoints is that the risk of dying from another cause than experiencing the endpoint of interest is generally much higher in older patients than in younger patients[4;5]. These so-called competing events are of major importance in the design and analyses of studies with older patients[5], as a patient who dies from another cause, can obviously not experience the event of interest anymore. This topic is especially important in geriatric oncology research, as a large proportion of older cancer patients will die from non-cancer related causes before reaching the endpoint of interest[5]. For example, around 70% of breast cancer patients aged 75 years or older who die, die from another cause than breast cancer[6].

There are several statistical methods that are frequently used for time to event analyses, such as the Cox Proportional Hazards Model and the Fine & Gray model. These methods deal with competing events in different ways[2]. It is likely that the choice of model can strongly influence the interpretation of the outcome, especially in populations with a high frequency of competing events[3;4]. Several studies have described the methodology of dealing with competing risks in detail, but these methodological papers may be difficult to interpret in clinical research. Therefore, the aim of this study is to present clinical examples of research questions in a population-based cohort of older breast cancer patients with a high frequency of competing events, and to discuss the implications of choosing different methods for the interpretation of the results. In addition, this paper will give recommendations in choosing the type of analyses for specific research questions.

Theoretical framework

First, we will provide some background information on methods that can be used to calculated absolute risks (i.e. cumulative incidences), and models that can be used to model the effect of variables on the outcome.

Estimating absolute risks

The Kaplan-Meier method is a commonly used method to estimate survival probabilities over

time. It can deal with censored follow-up times, i.e. it can handle situations where the exact time of death is not known because patients drop out of the study, or are still alive at the end of follow-up. One important assumption of the Kaplan-Meier method is independent censoring: at any time patients with censored survival times have the same survival prognosis as patients who are still in the study[7].

Kaplan Meier curves are often used to calculate survival probabilities for a specific cause of death. Patients who die of other causes are censored. Clearly, the assumption that censored patients have the same prognosis as those who are still followed is invalid, since patients who die of other causes have a probability of zero to reach the cause of interest. This means that estimated survival probabilities of the Kaplan-Meier method are no longer correct. Hence, the Kaplan-Meier method does not estimate the actual survival probability, but estimates what would have been observed if dying from other causes is not possible.

Alternatively, the Cumulative Incidence Competing Risks (CICR) method[2;3] assumes that patients who experienced a competing event are no longer at risk for the endpoint of interest[8]. This approach estimates the actual probabilities of reaching different endpoints (cumulative incidences). At each time point, the sum of all the cumulative incidences will be equal to the total probability to reach an endpoint before that time.

Hazard functions

There are different ways to assess the association between certain variables and the outcome with the possibility to adjust for confounding factors. The most commonly used methods are Cox proportional hazards models and Fine and Gray models. In order to understand the difference between these two models we first have to introduce the concept of the hazard function. Roughly speaking, hazard functions are event rates which vary over time. An intuitive explanation of the hazard can be given in the situation when time is discrete. In this instance, the hazard at a certain time is the probability to die at that time point, in those patients who are still alive[2] . In absence of competing risk there is a one to one mathematical relation between the hazard function and the survival function.

Cox proportional hazards model

The Cox proportional hazards model assesses the effects of variables on the hazard function. In the Cox proportional hazards model, hazard functions for different values of the prognostic variable are assumed to be proportional over time, and the parameters of the models can be interpreted as hazard ratios (HRs). In absence of competing risk, a HR above one implies smaller survival probabilities for the exposed group compared to the unexposed group.

In a similar way, cause specific hazard functions can be defined. Cause specific hazards are similar to cause specific mortality rates over small time periods. Effects of prognostic factors on cause specific hazards can be assessed using the Cox proportional hazard model, where subjects who die of other causes are censored. However, a HR above one no longer implies that subjects with the risk factor are truly more likely to experience the specific event, because subjects can die of other causes before they are able to reach this event. If the hazards for dying from other causes are much larger and the prognostic factor also affects these hazards, it could happen that actually less people reach the cause of interest. For example, smoking increases the hazard to develop dementia, but only few smokers will actually develop dementia, because of the competing effects of death due to cancer or cardiovascular diseases. As a result it could happen that actually less smokers than non-smokers will experience dementia (i.e the cumulative incidence of dementia is lower in the smoking group) , even though the cause specific HR for the effect of smoking on dementia may actually be larger than 1. Assuming that there is a causal relation between smoking and dementia, this relation can be found by the Cox Regression model (i.e the HR is higher than one for smokers), while the cumulative incidence of dementia is in fact lower in smokers due to competing causes of death.

Fine and Gray model

The Fine and Gray model[9] links the effect of risk factors directly to the cause specific cumulative incidences of death. In our smoking-dementia example, the Fine and Gray model considers the direct effect of smoking on the cumulative incidence of dementia (which was lower for smokers, due to the competing risks). The effect of risk factors are expressed in "subdistribution hazard ratios" (SHR), where the subdistribution hazard function has a one to one relation with the cause specific cumulative distribution function. An intuitive interpretation of this SHR is difficult but readers should remind that a SHR above one corresponds to higher cause specific event probabilities. In our dementia example, the Fine and Gray model will yield a SHR below 1, as it directly models the cumulative incidence of developing dementia in both subgroups, resulting in a lower risk of dementia for smokers. For a more detailed theoretical background of these models, we refer to the paper of Putter et al.[3]

Table 1 *Patient and tumour characteristics of patients in the FOCUS-cohort*

	N	%
Age		
65-74 years	1,425	50.8
≥75 years	1,380	49.2
Stage		
I	1,058	37.7
II	1,430	51.0
III	317	11.3
Grade		
1	385	13.7
2	906	32.3
3	670	23.9
missing	844	30.1
Morphology		
Ductal	2,074	73.9
Lobular	328	11.7
Mixed/other	403	14.4
Number of comorbid diseases		
0	694	24.7
1	656	23.4
2 or more	1,455	51.9

Clinical examples

For the examples in this paper, we used data from the population-based FOCUS-cohort (Female breast cancer in the elderly; Optimizing Clinical guidelines USing clinico-pathological & molecular data). This cohort comprises all incident breast cancer patients aged 65 years or older, who were diagnosed in the geographically defined Comprehensive Cancer Center Region West in The Netherlands between January 1997 and December 2004 (N=3,672). Trained personnel reviewed the charts of these patients, and collected information on specific treatments, comorbidity according to the ICD-10 classification[10], adverse events, geriatric parameters, and recurrences.

For the examples that are used below, only patients with non-metastatic, invasive breast cancer, who received primary surgery were included. The endpoint of interest was breast cancer recurrence, defined as any local recurrence (skin or in-breast), regional recurrence (axillary or supraclavicular lymph nodes) or distant metastasis. The competing event was defined as death due to any cause without breast cancer recurrence. Censoring only occurred due to end of follow-up or immigration, the latter being very rare in our cohort.

Overall, 2,805 patient were included in the analyses. Patient and tumour characteristics are briefly described in Table 1. Median follow-up time was 5.6 years, ranging from 0 to 14.2 years. Overall, 478 (17%) developed a breast cancer recurrence. The prevalence of competing events (death without recurrence) was 36% (N=1,015). The risk of competing events increased with age, from 19% in patients aged <75 years, to 54% in patients aged ≥75 years.

Example 1: Psychiatric disorders in association with breast cancer recurrence

Recently, we assessed the association between concomitant disease and breast cancer recurrence[11], as it has been suggested that concomitant diseases can interact with tumour growth as well as certain treatments[12]. Hence, the research question that we aimed to study was of an *etiological* nature. One of the concomitant diseases that we assessed were psychiatric disorders, defined according to the ICD10-classification[10]. We will now discuss several models that can be used to study the association between psychiatric disease and breast cancer recurrence. For simplicity reasons, we will present univariate analyses only.

Overall, 256 patients in the FOCUS cohort had a psychiatric disorder. Of all patients with a psychiatric disorder, 29 (11%) developed a breast cancer recurrence during follow-up. Among the 2,549 patients without psychiatric disorders, 449 (18%) developed a recurrence. Among patients with psychiatric disorders, 150 (59%) competing events occurred, as compared to 865 (34%) among patients without psychiatric disorders.

Table 2 *Cumulative incidence of recurrence in relation to psychiatric disorders*

	KM	CICR	HR*	95% CI	SHR**	95% CI
No psychiatric disorders	24%	20%	1 (ref)		1 (ref)	
Psychiatric disorders	18%	12%	0.78	(0.53-1.13)	0.61	(0.42-0.9)

First, we assessed the association of psychiatric disease with breast cancer recurrence using the Kaplan Meier method. The 10-year cumulative incidence of breast cancer recurrence as calculated by the Kaplan Meier method in patients without a psychiatric disorder was 24%, compared to 18% among patients with a psychiatric disorder (Table 2). Second, we used the CICR method to assess cumulative incidence of recurrence, which resulted in a 10-year cumulative incidence of recurrence of 20% and 12%, respectively for patients without and with a psychiatric disorder. This shows that the Kaplan Meier method overestimates the cumulative incidences.
As shown in Table 2, the HR for having a recurrence for patients with a psychiatric disorder was 0.78 (95% confidence interval (CI) 0.53-1.13) compared to patients without a psychiatric disorder, calculated by a unadjusted Cox regression analysis. This implies that there is no significant difference in the hazard on recurrences between patients with and without psychiatric disorders. The hazards were proportional over time (tested using Schoenfeld residuals p=0.27).

Patients with a psychiatric disorder had a higher probability to die of any cause (HR 1.6, 95% C.I. 1.4-1.8, p<0.001, compared to patients without psychiatric diseases). In Fine & Gray regression analysis, the SHR was 0.61 (95% CI 0.42-0.90) for patients with a psychiatric disorder, as compared to patients without psychiatric disorders. This implies that the probability of recurrence was estimated to be lower for patients with psychiatric disease when the Fine & Gray model was used, compared with the Cox Regression Model. In this example, the Fine & Gray model, in contrast with the Cox model, even yielded a result which was statistically significant.

Example 2: *Prediction* of breast cancer recurrence in older patients

Currently, interest is in the prediction of the risk of breast cancer recurrence and breast cancer mortality, in order to estimate which patients are at high risk and should receive additional treatments. Most of currently available models were developed in generally young populations, and were not validated in older populations[13].We recently showed that the online Adjuvant! program, which is widely implemented in daily clinical practice, does not accurately predict breast cancer recurrence in older patients[14]. Therefore, one of the aims of the FOCUS-study is to develop a new prediction tool that can be used to estimate breast cancer recurrence in older patients. Hence, for this study, we were interested in *predictors* of breast cancer recurrence, and in calculating the absolute risk of recurrence.

In a recent review, it was shown that both tumour size and nodal status are the most incorporated variables in prediction models for breast cancer prognosis.[15] Therefore, for this example, we assessed the predictive value of tumour stage according to the Tumour-Node-Metastasis (TNM) classification. Cumulative incidences of recurrence and competing events by stage, calculated by the Kaplan-Meier method and the CICR function are presented in Figure 1a and 1b respectively. As shown in Figure 1a, towards the end of follow-up, the cumulative mortality and the cumulative incidence of recurrence as calculated by the Kaplan-Meier method added up to estimates higher than 100%, as the lines in the figure cross. In contrast, the sum of the estimates of mortality

Figure 1A *Recurrence-free period and cumulative incidence of competing events using the Kaplan Meier method*

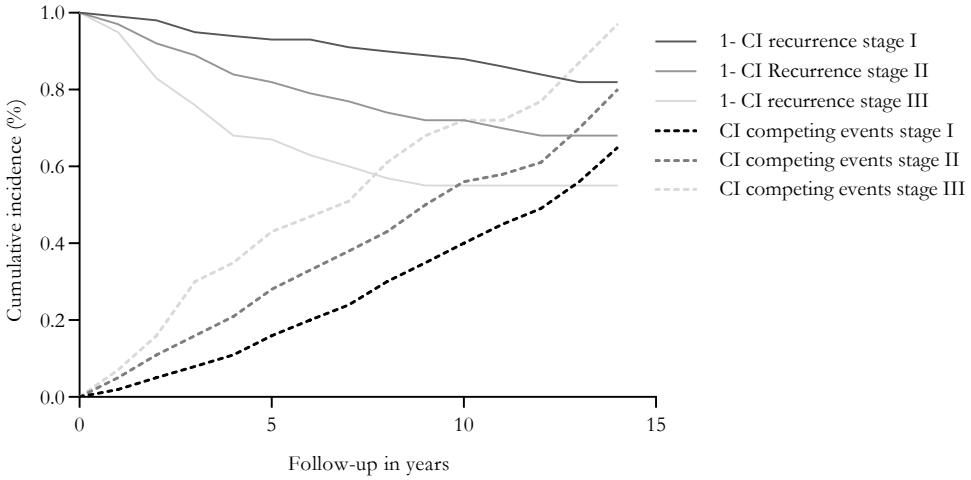
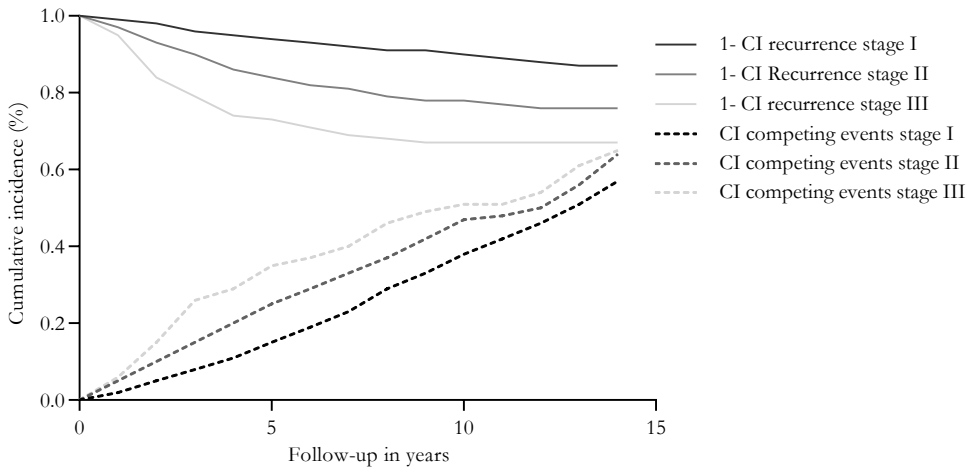


Figure 1B *Recurrence-free period and cumulative incidence of competing events using the CICR method*



CI=Cumulative incidence; competing events are all deaths without recurrence

and recurrence did not exceed 100% when the CICR method is used (Figure 1b). Clearly, the Kaplan-Meier method overestimated the cumulative incidence of recurrence and the cumulative incidence of competing events.

In order to further demonstrate the impact of competing events, we stratified patients into two age-groups: <75 years and ≥75 years (Table 3b and 3c). In both age groups, tumour stage was predictive for breast cancer recurrence, as can be expected. However, in patients aged <75 years, the prevalence of competing events was 19% during follow-up, which is much lower than in patients aged ≥75 years, where 54% of all patients died without a recurrence. Among patients aged <75 years, the lower incidence of competing events as compared to patients >75 years resulted in relatively small differences in outcomes between the Kaplan-Meier method and the CICR method, while in the patients >75 years, the Kaplan-Meier method more strongly overestimated the risk of recurrence.

Cox Regression analyses resulted in a strongly increased risk of recurrence with increasing tumour stage (HR 5.42, 95% CI 4.08-7.21 for stage III versus stage I, Table 3). Although the difference between the tumour stages remained statistically significant in the Fine & Gray analysis, the Fine & Gray analysis attenuated the effect estimates. For predictive research, we are interested in the direct effect on the cumulative incidence and therefore Fine & Gray analyses provide more valid effect estimates. As shown in Table 3, the differences between estimates that

Table 3 *Cumulative incidence of recurrence by stage*

10 year cumulative recurrence for all patients						
	KM	CI	HR*	95% CI	SHR**	95% CI
Stage I	12%	10%	1 (reference)		1 (reference)	
Stage II	28%	22%	2.72	(2.15-3.44)	2.43	(1.93-3.07)
Stage III	45%	33%	5.42	4.08	4.10	3.08
3a. 10 years cumulative recurrence for patients aged <75 years						
	KM	CI	HR*	95% CI	SHR**	95% CI
Stage I	12%	11%	1 (reference)		1 (reference)	
Stage II	27%	24%	2.68	(1.98-3.63)	2.59	(1.91-3.50)
Stage III	55%	46%	6.62	(4.51-9.72)	5.72	(3.89-8.41)
3b. 10 years cumulative recurrence for patients aged ≥ 75 years						
	KM	CI	HR*	95% CI	SHR**	95% CI
Stage I	12%	9%	1 (reference)		1 (reference)	
Stage II	29%	20%	2.56	(1.75-3.75)	2.32	(1.59-3.40)
Stage III	35%	26%	4.31	(2.77-6.71)	3.35	(2.15-5.23)

**derived from univariable Cox regression analysis **derived from Fine & Gray analyses*

are calculated in Cox Regression analyses and Fine & Gray analyses become larger when the frequency of competing events increases.

Reflection

Our results show that in populations with a high frequency of competing events, it is important to consider which methods are the most appropriate to deal with cause-specific endpoints. The Kaplan Meier method should never be used to estimate cause specific survival curves since it overestimates the absolute risk of the event of interest. The CICR method appropriately deals with competing risks. When assessing relative effect sizes in *etiologic* research, the Cox Proportional Hazards model is most appropriate. In contrasts, for absolute risk estimates in *predictive* research, the Fine & Gray Model should be used in populations with a high frequency of competing events.

The main strength of this paper is that the examples were performed using a real cohort of patients with a high prevalence of competing risk. By presenting the results of several methods in different research questions, we were able to demonstrate the effects of the choice of a certain method in different settings. Of course, this study also has its limitations. First, it must be noted that the recurrence rate that was registered in the cohort may have been underestimated, as older patients may be less adherent to follow-up schemes. This may have influenced our analyses, especially if there was selective non-adherence to follow-up schemes. In addition, 10-year follow-up for recurrence was not complete for the whole cohort, but this mostly applied to the most recent years of the cohort, and it is unlikely that this has influenced our results as it has been shown that outcome of older patients has not changed in recent years[17].

With the results of our current study, we want to highlight the difference between *etiological* and *predictive* research questions in the comparison between the Cox Proportional Hazards model and the Fine & Gray model. In Example 1, the Fine & Gray model yielded rather strange results from an etiological point of view as is suggests that psychiatric disorders are protective for recurrence. It is very unlikely that there is some biological mechanism in which psychiatric disorders are protective for breast cancer recurrence. More likely, our finding can be explained by the fact that the Fine & Gray analysis incorporates the competing risk of death which influences the cumulative incidences of recurrence This makes sense, since patients with psychiatric disorders (especially dementia) have an increased risk of dying compared to patients without psychiatric disorders and patients who have died, cannot experience a breast cancer recurrence anymore. In contrast, the Cox Regression Model considers the effect on the cause specific hazards, i.e on the

instantaneous risk of recurrence for patients who are still at risk for the event at a certain time-point, and this is what we are interested in in this research question.

Therefore, in etiologic research questions, the Cox Regression model is often the most appropriate method. In contrast, for *predictive* studies, methods that incorporate competing events such as Fine & Gray competing risk regression are more appropriate. In prediction, we are generally interested in calculating absolute risks rather than relative risks, and in this case it is important to consider that patients with a large risk of experiencing a competing event are unlikely to develop a breast cancer recurrence.

In conclusion, in studies with cause-specific endpoints in populations with a high frequency of competing events, researchers should carefully choose the most appropriate statistical method in order to prevent incorrect interpretation of study results.

References

(1) de Glas NA, Hamaker ME, Kiderlen M et al. Choosing relevant endpoints for older breast cancer patients in clinical trials: an overview of all current clinical trials on breast cancer treatment. *Breast Cancer Res Treat* 2014;146(3):591-7.

(2) Andersen PK, Geskus RB, de WT, Putter H. Competing risks in epidemiology: possibilities and pitfalls. *Int J Epidemiol* 2012;41:861-870.

(3) Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. *Stat Med* 2007;26:2389-2430.

(4) Berry SD, Ngo L, Samelson EJ, Kiel DP. Competing risk of death: an important consideration in studies of older adults. *J Am Geriatr Soc* 2010;58:783-787.

(5) Pallis AG, Ring A, Fortpied C et al. EORTC workshop on clinical trial methodology in older individuals with a diagnosis of solid tumors. *Ann Oncol* 2011;22:1922-1926.

(6) van de Water W, Markopoulos C, van de Velde CJ et al. Association between age at diagnosis and disease-specific mortality among postmenopausal women with hormone receptor-positive breast cancer. *JAMA* 2012;307:590-597.

(7) Bland JM, Altman DG. Survival probabilities (the Kaplan-Meier method). *BMJ* 1998;317:1572.

(8) Verduijn M, Grootendorst DC, Dekker FW, Jager KJ, le CS. The analysis of competing events like cause-specific mortality--beware of the Kaplan-Meier method. *Nephrol Dial Transplant* 2011;26:56-61.

(9) Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association* 1999;94:496-509.

(10) World Health Organisation . The ICD-10 Classification of Mental and Behavioural Disorders. 10[5], 1-267. 1992.

(11) Kiderlen M, de Glas NA, Bastiaannet E et al. Impact of comorbidity on outcome of older breast cancer patients: a FOCUS cohort study. *Breast Cancer Res Treat* 2014;145(1):185-92.

(12) Patnaik JL, Byers T, Diguiseppi C, Denberg TD, Dabelea D. The influence of comorbidities on overall survival among older women diagnosed with breast cancer. *J Natl Cancer Inst* 2011;103:1101-1111.

(13) Engelhardt EG, Garvelink MM, de Haes JH et al. Predicting and Communicating the Risk of Recurrence and Death in Women With Early-Stage Breast Cancer: A Systematic Review of Risk Prediction Models. *J Clin Oncol* 2013;2014;32(3): 238-50.

(14) de Glas NA, van de Water W, Engelhardt EG et al. Validity of Adjuvant! Online program in older patients with breast cancer: a population-based study. *Lancet Oncol* 2014;15:722-729.

(15) Engelhardt EG, Garvelink MM, de Haes JH et al. Predicting and communicating the risk of recurrence and death in women with early-stage breast cancer: a systematic review of risk prediction models. *J Clin Oncol* 2014;32:238-250.

(16) Fleming I.D., Cooper JS, Henson D.E., Hutter RVP, Kennedy B.J., et al. *AJCC Cancer Staging Manual, fifth edition*. 5th ed. Philadelphia, PA: Lippincott-Raven, 1997.

(17) de Glas NA, Jonker JM, Bastiaannet E et al. Impact of omission of surgery on survival of older patients with breast cancer. *Br J Surg* 2014;146(3):591-7.

(18) Wildiers H, Mauer M, Pallis A et al. End points and trial design in geriatric oncology research: a joint European organisation for research and treatment of cancer-alliance for clinical trials in oncology-international society of geriatric oncology position article. *J Clin Oncol* 2013;31:3711-3718.

(19) Dignam JJ, Huang L, Ries L, Reichman M, Mariotto A, Feuer E. Estimating breast cancer-specific and other-cause mortality in clinical trial and population-based cancer registry cohorts. *Cancer* 2009;115:5272-5283.

(20) Goldoni CA, Bonora K, Ciatto S et al. Misclassification of breast cancer as cause of death in a service screening area. *Cancer Causes Control* 2009;20:533-538.

(21) van de Water W, Kiderlen M, Bastiaannet E et al. External validity of a trial comprised of elderly patients with hormone receptor-positive breast cancer. *J Natl Cancer Inst* 2014;106:dju051.

(22) Blakely T, Soeberg M, Carter K, Costilla R, Atkinson J, Sarfati D. Bias in relative survival methods when using incorrect life-tables: lung and bladder cancer by smoking status and ethnicity in New Zealand *Int J Cancer* 2012;131:E974-E982.