

Aanmeldingsformulier beleidsAiO 2005 (max. 3 pages, Times New Roman minimal 10 pts)				
1. Title: Models for the Estimation of Causal Treatment Effects with Observational Data				
2. Project leader(s): R.B. Geskus, PhD Department of Clinical Epidemiology & Biostatistics, AMC				
3. Present scientific members of the research team:				
name	Function	ODP	financial support	Fte
RB Geskus	UD	AmCOGG 15	AMC	0.1
AH Zwinderman	Professor	AmCOGG 15	AMC	0.05
J Korevaar	UD	AmCOGG 15	AMC	0.05
F de Wolf	director		HIV monitoring foundation	0.05
M Prins		INFE 07	GGD Amsterdam	0.05
4. Summary Individuals with an elevated disease or death risk are often monitored over time. At each new visit, an individual's health state is determined by measuring variables, called markers, which reflect disease progression. Individuals with worse disease prognosis are more likely to initiate or change treatment. Information obtained from such observational databases can be used to model the effect of treatment on disease progression. Clinical trials are the standard setting in which to investigate treatment effects, but they are often hampered by ethical, financial or logistic constraints. The lack of randomisation of treatment administration in observational databases invalidates the use of standard statistical methods. Dedicated methods that deal with non-randomised settings are required. These methods have been under development over the last decade, but are still little used by the research community. The aim of the project is to investigate and extend existing methods and apply them to three observational databases. One database has monitoring information on dialysis patients who at some point may switch treatment from peritoneal dialysis to hemodialysis. A second data set concerns the effect of borrowing used injecting equipment on progression to AIDS among injecting drug users. The third database has information on treatment effects and AIDS progression in HIV-infected individuals. We will obtain better methods to quantify the causal effect of treatment in observational databases of patients who are monitored over time. These methods will be applicable to a wide range of clinical studies as performed within AmCOGG, using the software that will be developed as part of the project.				
5. Project:				
a) purpose The aim of the project is to investigate and extend existing models that allow for the estimation of the causal effect of treatment on subsequent disease progression in non-randomised settings in which patients are monitored over time.				
b) background and previous research Estimation of the causal effect of treatment on some outcome (mortality, disease, marker development) is best performed in randomised experiments like clinical trials. Often, the practical implementation of such trials is hampered by ethical, financial or logistic constraints. Observational databases that contain information on patients who have been monitored over time are the only or the more extensive source of information on the treatment effect. However, confounding is a major concern since the randomisation property is lost: individuals who are in worse clinical condition, defined by the value of one or more markers, are more likely to start or change treatment. The estimated causal effect of a beneficial treatment will be biased towards the null, or may even be reversed, unless some correction is made for clinical condition. Simply including the marker values as cofactors in the analysis is also problematic, since treatment also influences the subsequent marker development. The estimated treatment effect will be attenuated, since the markers are intermediate in the causal pathway from the treatment to the outcome of interest. A variable that (1) is a time-dependent risk factor for the outcome of interest (a marker), and (2) predicts subsequent treatment allocation, is called a time-dependent confounder. Typically, (3) treatment also influences subsequent marker development. It has been shown that the associations estimated from standard models for the effect of treatment, e.g. by modelling the event risk as a function of past exposure and covariate history in a time-dependent Cox proportional hazards model, gives biased results on the causal effect of treatment under the presence of a time-dependent confounder which is also influenced by past treatment history (Robins, 1997). It is widely acknowledged that association is not causation. One can say that there is a causal effect of some exposure or				

intervention on the outcome if the outcome depends on whether the individual received the intervention or not. However, only one situation is observed per the individual: either the intervention was present, or the intervention was absent. The other situation is *counterfactual* and it is unknown what would have happened in that situation. Fortunately, at the population level, estimation of causal effects may still be possible. In case of a randomised experiment, this is a consequence of the exchangeability property (Hernan, 2004). However, the counterfactual theory also allows for the estimation of causal effects in observational settings in which the probability to receive the intervention is determined by clinical condition, as long as there is conditional exchangeability: all relevant clinical conditions need to be included in the model (no unmeasured confounding). Since this cannot be verified from the data, one has to rely on expert opinion.

Robins extended the counterfactual model to the estimation of causal effects of time varying exposures. Structural nested failure time models (SNM; Robins, 1998) and marginal structural models (MSM; Robins, 1998) can be used to estimate the causal effect of treatment on the outcome of interest. Recently, an extension of MSMs has been developed, called history-adjusted marginal structural models (HA-MSM; van der Laan and Petersen, 2004). HA-MSMs allow for the investigation of the causal effect of treatment based on a shorter, user-specified, exposure history than MSMs and allow the causal effect of treatment to depend on the value of time-dependent covariates. They can be used to find optimal *dynamic* treatment regimes, which are regimes that can change over time in response to the development of time-dependent covariates (i.e. markers). These models have most often been used in AIDS research, in which CD4 count is a time-dependent confounder for the effect of treatment on future CD4 development (Hernan *et al.*, 2002) and progression to AIDS (Cole *et al.*, 2003). Using observational databases, causal treatment effects were estimated and compared with the results from traditional approaches, showing that the latter provide highly biased causal effect estimates.

c) global approach

Although the importance of drawing proper causal inferences is widely acknowledged, so far little use has been made of the methods described in b). Lack of expertise in the use of these methods is probably the main reason for this. Our project aims to develop this expertise at the AMC and applies the methods to three different data sets. Emphasis will be on the MSM and its extension HA-MSM. These methods have the advantage that, in principle, standard models like the Cox proportional hazards model can be used. The difference is in the use of a weighting factor, resulting in “inverse of probability of treatment weighted” (IPTW) and “inverse of probability of censoring” estimators, which are related to propensity scores. The beleidsaio will become familiar with these methods and the difficulties involved in the use of these methods: the data manipulation needed before the methods can be applied, the proper choice of model in the calculation of the weights, and the choice of the correct model for the causal treatment effect. Cumulative treatment effects have most often been used, but HA-MSMs also allow for the modelling of temporary treatment effects and interaction with time-dependent covariates. Sensitivity of IPTW estimators to unmeasured confounding (Brumback *et al.*, 2004) will also be investigated. Furthermore, models that jointly describe marker development and time to some event (Geskus *et al.*, 2005) will be extended by including a causal treatment effect.

The data sets used are now briefly described.

1. The Netherlands Cooperative Study on the Adequacy of Dialysis (NECOSAD) is a large multicenter prospective observational cohort study in which patients with end stage renal disease are monitored from start of hemodialysis (HD) or peritoneal dialysis (PD) until renal transplantation or death. In Termorshuizen *et al.*, 2003, mortality rates for patients under HD were compared with those for patients under PD. Their approach did not allow for evaluation of the effect of treatment switch from PD to HD. The three conditions that make the standard approaches invalid are satisfied. Residual renal function has been shown to be an important marker for patient survival in both groups of patients (1). Since peritoneal dialysis is thought to be less effective in patients with low residual renal function, such patients are more likely to switch to hemodialysis (2). In the NECOSAD database, more than 25% of the patients on peritoneal dialysis switched to hemodialysis during follow-up. Moreover, the development of residual renal function differs by dialysis group (3). We will use the above methods in order to evaluate the causal effect of the switch to HD.
2. The Amsterdam cohort study on drug users was started in 1985 in order to assess the prevalence and incidence of HIV infection, to determine risk factors for HIV infection and to evaluate the natural history of HIV infection among drug users. One study found a significantly slower AIDS progression for drug users who more often had borrowed used injecting equipment before entry into the study or seroconversion (Mientjes *et al.*, 1993). CD4 count influences AIDS progression (1), but may also determine the exposure variable of interest: the frequency of borrowing (2). Then, CD4 count is a time-dependent confounder for the effect of borrowing on AIDS development. But borrowing may also influence subsequent CD4 count (3). Hence, when evaluating the causal effect of borrowing on AIDS progression, the above methods need to be used.

3. Since 1996, highly active anti-retroviral therapy (HAART) has greatly improved prognosis of HIV infected individuals. However, due to the many side effects of HAART, the choice of the optimal moment of initiation of therapy and the optimal strategy after HAART initiation are important. Considerations for initiation of HAART are based on the possibility of the immune system to recover after administration of HAART, the side effects to be expected, and on what will happen if persons are not treated (Geskus *et al.*, 2003). Data from the HIV monitoring foundation (<http://www.hiv-monitoring.nl>), which collects data of all individuals on HAART in the Netherlands, will be used to evaluate the causal effect of HAART on CD4 development, time to AIDS and time to treatment failure.

d) cited literature (max. 15 incl. own publications, Times New Roman 8 pt.)

Brumback BA, Hernan MA, Haneuse SJPA et al. Sensitivity analyses for unmeasured confounding assuming a marginal structural model for repeated measures. *Stat. Med.* **2004**, 23: 749-767.

Cole SR, Hernan MA, Robins JM, et al. Effect of highly active antiretroviral therapy on time to AIDS or death using marginal structural models. *Am J Epidemiol* **2003**, 158: 687-694.

Geskus RB, Miedema FA, Goudsmit J et al. Prediction of residual time to AIDS and death based on markers and cofactors. *JAIDS* **2003**, 32: 514-521.

Geskus RB, Meyer L, Hubert J-B et al. Causal pathways of the effects of age and the CCR5-Δ32, CCR2-64I and SDF-1 3'A alleles on AIDS development. *JAIDS* **2005**, 39: 321-326.

Hernan MA, Brumback BA, Robins JM. Estimating the causal effect of zidovudine on CD4 count with a marginal structural model for repeated measures. *Stat Med* **2002**, 23: 749-767.

Hernan MA. A definition of causal effect for epidemiological research. *J Epidemiol Community Health* **2004**; 58: 265-271.

Mientjes GHC, van Ameijden EJC, van den Hoek EJC et al. Progression of HIV infection among injecting drug users: indications for a lower rate of progression among those who have frequently borrowed injecting equipment. *AIDS* **1993**; 7: 1363-1370.

Robins JM. Causal inference from complex longitudinal data. In: Barkane M, ed. *Latent Variable Modeling and Applications to Causality: Lecture Notes in Statistics* 120. New York: Springer-Verlag, **1997**: 69-117.

Robins JM. Marginal structural models. 1997 Proc. of the Section on Bayesian Statistical Science. Alexandria, Virginia: Am. Stat. Assoc. **1998**: 1-10.

Robins JM. Structural nested failure time models. In: Andersen PK, Keiding N, section eds. *Survival Analysis*. In: Armitage P, Colton T, eds. *The Encyclopedia of Biostatistics*. Chichester, UK: John Wiley and Sons, **1998**: 4372-4389.

Termorshuizen F, Korevaar JC, Dekker FW, et al. Hemodialysis and peritoneal dialysis: comparison of adjusted mortality rates according the duration of dialysis: analysis of the Netherlands cooperative study on the adequacy of dialysis 2. *J Am Soc Nephrol* **2003**; 14: 2851-2860.

Van der Laan MJ and Petersen ML. History-adjusted marginal structural models and statically-optimal dynamic treatment regimes (<http://www.bepress.com/ucbbiostat/paper158/>, published online on September 24, **2004**)

6. Planning of PhD research project:

a) first year

The appointed individual needs to have a decent background in statistics, but no knowledge of causal models is required. This will be acquired during the first year. Practical experience will be obtained in using MSM models: data from the second study will be used in order to analyse the effect of the borrowing of used needles on AIDS risk. The results from this study, which are expected to be different from the previous results in Mientjes *et al.*, will be written down in a scientific paper. MJ van der Laan will be invited for a two-week visit to our department in order to give an introduction into HA-MSMs.

b) next 3 years (globally)

The beleidsaio will further study the theory on HA-MSMs. HA-MSMs will be used in the analyses of studies 1 and 3. In both studies, the causal effects of treatment initiation and treatment change on marker development as well as on mortality and other end points will be investigated. HA-MSMs will also be used in order to model different causally relevant exposure histories.

RB Geskus has been a member of the writing committee of a proposal to use causal models to estimate the effect of modern treatment regimens for HIV infected patients (NIH grant "International Epidemiologic Databases to Evaluate AIDS (IEDEA)": <http://grants.nih.gov/grants/guide/rfa-files/RFA-AI-05-014.html>). If this IEDEA proposal is accepted, the beleidsaio will be involved in this project.

Results from these studies will be written down in scientific papers and will be presented at international conferences.

Date: 27-07-2005

Signature(s) project leader(s):