Appendix 1. Development of the 'Optimal fit' model

The trial data were regarded as a single prospective cohort with variable follow-up time. The occurrence of the primary combined end point of myocardial infarction, stroke, arterial revascularization, hospitalization for unstable angina, or death from cardiovascular causes was modelled using Cox proportional hazard regression. To minimize over-fitting, candidate predictors for the outcome and transformations of variables were based on existing prediction models (the Framingham risk score and the Reynolds risk score).[1-3] The starting model, therefore, contained the following variables: rosuvastatin treatment (yes/no), age, gender, current smoking status, log transformed systolic blood pressure (SBP), log transformed high density lipoprotein cholesterol (HDLc), log transformed total cholesterol (TC), log transformed high-sensitivity C-reactive protein, use of blood pressure lowering medication, and family history of premature coronary heart disease. Parallel to the Framingham risk score, interactions between age and gender and between log transformed SBP and blood pressure lowering medication were also considered.[1] Continuous predictors were truncated at the 1st and 99th percentile to limit the effect of outliers. The model was backward selected on the basis of the Akaike's Information Criterion (AIC). To minimize over-fitting, the simplest model was chosen if the model AIC increased by less than 2 after deletion of a factor (Occam's razor).[4] Interaction of treatment effect with any remaining covariates was considered after initial backward selection of the starting model. The resulting 'Optimal fit'-model was used to predict each patient's 2 year risk for cardiovascular events with and without rosuvastatin treatment. Based on the assumption of constant hazard and, thus, exponential risk over time, the 2 year risk estimates were extrapolated to 10 years. The predicted absolute risk reduction achieved by rosuvastatin treatment at 10 years (10-year treatment effect) was defined as the difference between the two risk predictions (Box 1).

1 National Cholesterol Education Program. Executive summary of the third report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). *JAMA* 2001;285:2486-97.

2 Ridker PM, Buring JE, Rifai N, Cook NR. Development and validation of improved algorithms for the assessment of global cardiovascular risk in women: the Reynolds risk score. *JAMA* 2007;297:611-9.

3 Ridker PM, Paynter NP, Rifai N, Gaziano JM, Cook NR. C-reactive protein and parental history improve global cardiovascular risk prediction: the Reynolds risk score for men. *Circulation* 2008;118:2243-51.

4 Burnham KP, Anderson DR. Model selection and multimodel inference: a practical information-theoretic approach. 2nd ed. Springer, 2002.