# **Table S1. Average bias, variance, and MSE of point estimators of hazard ratio in the M+ subgroup: additional scenarios.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|   |   |   | *E* = 250 |  | *E* = 350 |
| HR+ | HR- | *p*  | *Q* |   | Standard | Naive correction | *a*1(Zint) | *a*2(Zint) |  | *Q* |   | Standard | Naive correction | *a*1(Zint) | *a*2(Zint) |
| 0.5 | 0.7 | 0.3 | 0.473 | Bias | -0.079 | -0.036 | 0.022 | 0.045 |  | 0.566 | Bias | -0.055 | -0.020 | 0.027 | 0.043 |
|  |  |  |  | Var | 0.007 | 0.010 | 0.018 | 0.026 |  |  | Var | 0.005 | 0.008 | 0.014 | 0.019 |
|  |  |  |  | MSE | 0.013 | 0.012 | 0.018 | 0.028 |  |  | MSE | 0.008 | 0.009 | 0.014 | 0.021 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  | 0.6 | 0.507 | Bias | -0.039 | -0.017 | 0.012 | 0.021 |  | 0.601 | Bias | -0.026 | -0.008 | 0.016 | 0.022 |
|  |  |  |  | Var | 0.005 | 0.006 | 0.008 | 0.010 |  |  | Var | 0.004 | 0.005 | 0.006 | 0.008 |
|  |  |  |  | MSE | 0.007 | 0.007 | 0.009 | 0.011 |  |  | MSE | 0.004 | 0.005 | 0.007 | 0.008 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 0.5 | 0.8 | 0.3 | 0.665 | Bias | -0.050 | -0.012 | 0.041 | 0.057 |  | 0.763 | Bias | -0.033 | -0.004 | 0.035 | 0.044 |
|  |  |  |  | Var | 0.008 | 0.012 | 0.021 | 0.030 |  |  | Var | 0.007 | 0.010 | 0.017 | 0.022 |
|  |  |  |  | MSE | 0.011 | 0.013 | 0.023 | 0.033 |  |  | MSE | 0.008 | 0.010 | 0.018 | 0.024 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  | 0.6 | 0.708 | Bias | -0.033 | -0.004 | 0.035 | 0.044 |  | 0.809 | Bias | -0.012 | 0.002 | 0.021 | 0.024 |
|  |  |  |  | Var | 0.007 | 0.010 | 0.017 | 0.022 |  |  | Var | 0.004 | 0.005 | 0.007 | 0.008 |
|  |  |  |  | MSE | 0.008 | 0.010 | 0.018 | 0.024 |  |  | MSE | 0.004 | 0.005 | 0.008 | 0.009 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 0.5 | 1 | 0.3 | 0.892 | Bias | -0.016 | 0.013 | 0.051 | 0.057 |  | 0.948 | Bias | -0.004 | 0.013 | 0.036 | 0.036 |
|  |  |  |  | Var | 0.011 | 0.017 | 0.027 | 0.034 |  |  | Var | 0.01 | 0.013 | 0.02 | 0.023 |
|  |  |  |  | MSE | 0.011 | 0.017 | 0.029 | 0.037 |  |  | MSE | 0.01 | 0.013 | 0.021 | 0.024 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  | 0.6 | 0.922 | Bias | -0.007 | 0.006 | 0.022 | 0.023 |  | 0.967 | Bias | -0.001 | 0.006 | 0.015 | 0.014 |
|  |  |  |  | Var | 0.007 | 0.008 | 0.011 | 0.012 |  |  | Var | 0.005 | 0.006 | 0.007 | 0.008 |
|  |  |  |  | MSE | 0.007 | 0.008 | 0.011 | 0.012 |  |  | MSE | 0.005 | 0.006 | 0.008 | 0.008 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 0.5 | 1.2 | 0.3 | 0.968 | Bias | -0.001 | 0.017 | 0.041 | 0.041 |  | 0.992 | Bias | 0.002 | 0.010 | 0.020 | 0.018 |
|  |  |  |  | Var | 0.014 | 0.019 | 0.028 | 0.032 |  |  | Var | 0.011 | 0.013 | 0.017 | 0.018 |
|  |  |  |  | MSE | 0.014 | 0.019 | 0.030 | 0.034 |  |  | MSE | 0.011 | 0.013 | 0.018 | 0.018 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  | 0.6 | 0.981 | Bias | 0.000 | 0.008 | 0.016 | 0.015 |  | 0.996 | Bias | 0.002 | 0.005 | 0.008 | 0.007 |
|  |  |  |  | Var | 0.007 | 0.009 | 0.010 | 0.011 |  |  | Var | 0.005 | 0.006 | 0.006 | 0.006 |
|  |  |  |  | MSE | 0.007 | 0.009 | 0.011 | 0.011 |  |  | MSE | 0.005 | 0.006 | 0.006 | 0.006 |

HR+ (HR-): hazards ratio for M+ (M-) subgroup, *E*: the number of events at the final analysis,

*p*: observed prevalence of M+ subgroup in the study, *Q*: the empirical probability that the first interaction test is significant.

**Table S1 (continued)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |   |   | *E* = 250 |  | *E* = 350 |
| HR+ | HR- | *p*  | *Q* |   | Standard | Naive correction | *a*1(Zint) | *a*2(Zint) |  | *Q* |   | Standard | Naive correction | *a*1(Zint) | *a*2(Zint) |
| 0.6 | 0.7 | 0.3 | 0.247 | Bias | -0.137 | -0.084 | -0.011 | 0.022 |  | 0.263 | Bias | -0.110 | -0.065 | -0.003 | 0.024 |
|  |  |  |  | Var | 0.007 | 0.011 | 0.019 | 0.028 |  |  | Var | 0.006 | 0.008 | 0.014 | 0.020 |
|  |  |  |  | MSE | 0.026 | 0.018 | 0.019 | 0.029 |  |  | MSE | 0.018 | 0.013 | 0.014 | 0.021 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  | 0.6 | 0.259 | Bias | -0.072 | -0.043 | -0.003 | 0.013 |  | 0.283 | Bias | -0.060 | -0.036 | -0.003 | 0.009 |
|  |  |  |  | Var | 0.006 | 0.007 | 0.010 | 0.012 |  |  | Var | 0.005 | 0.006 | 0.007 | 0.009 |
|  |  |  |  | MSE | 0.011 | 0.009 | 0.010 | 0.012 |  |  | MSE | 0.008 | 0.007 | 0.007 | 0.009 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 0.6 | 0.8 | 0.3 | 0.404 | Bias | -0.1 | -0.049 | 0.023 | 0.051 |  | 0.494 | Bias | -0.075 | -0.033 | 0.025 | 0.046 |
|  |  |  |  | Var | 0.008 | 0.013 | 0.022 | 0.033 |  |  | Var | 0.007 | 0.01 | 0.017 | 0.024 |
|  |  |  |  | MSE | 0.018 | 0.015 | 0.023 | 0.035 |  |  | MSE | 0.012 | 0.011 | 0.018 | 0.027 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  | 0.6 | 0.428 | Bias | -0.051 | -0.023 | 0.014 | 0.027 |  | 0.517 | Bias | -0.038 | -0.016 | 0.014 | 0.023 |
|  |  |  |  | Var | 0.007 | 0.008 | 0.011 | 0.014 |  |  | Var | 0.005 | 0.006 | 0.008 | 0.01 |
|  |  |  |  | MSE | 0.009 | 0.009 | 0.011 | 0.014 |  |  | MSE | 0.006 | 0.006 | 0.008 | 0.011 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 0.6 | 1 | 0.3 | 0.710 | Bias | -0.050 | -0.007 | 0.054 | 0.070 |  | 0.812 | Bias | -0.027 | 0.005 | 0.048 | 0.058 |
|  |  |  |  | Var | 0.012 | 0.018 | 0.031 | 0.043 |  |  | Var | 0.010 | 0.015 | 0.025 | 0.032 |
|  |  |  |  | MSE | 0.014 | 0.018 | 0.034 | 0.048 |  |  | MSE | 0.011 | 0.015 | 0.027 | 0.036 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  | 0.6 | 0.755 | Bias | -0.023 | -0.001 | 0.027 | 0.033 |  | 0.858 | Bias | -0.011 | 0.005 | 0.025 | 0.027 |
|  |  |  |  | Var | 0.008 | 0.010 | 0.014 | 0.016 |  |  | Var | 0.006 | 0.008 | 0.010 | 0.011 |
|  |  |  |  | MSE | 0.009 | 0.010 | 0.015 | 0.018 |  |  | MSE | 0.006 | 0.008 | 0.011 | 0.012 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 0.6 | 1.2 | 0.3 | 0.886 | Bias | -0.020 | 0.013 | 0.057 | 0.064 |  | 0.972 | Bias | -0.004 | 0.016 | 0.042 | 0.042 |
|  |  |  |  | Var | 0.016 | 0.023 | 0.037 | 0.047 |  |  | Var | 0.013 | 0.017 | 0.026 | 0.030 |
|  |  |  |  | MSE | 0.016 | 0.023 | 0.041 | 0.051 |  |  | MSE | 0.013 | 0.018 | 0.028 | 0.032 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  | 0.6 | 0.918 | Bias | -0.004 | 0.011 | 0.030 | 0.031 |  | 0.972 | Bias | 0.001 | 0.009 | 0.020 | 0.019 |
|  |  |  |  | Var | 0.009 | 0.011 | 0.015 | 0.017 |  |  | Var | 0.007 | 0.008 | 0.010 | 0.011 |
|  |  |  |  | MSE | 0.009 | 0.012 | 0.016 | 0.018 |  |  | MSE | 0.007 | 0.008 | 0.011 | 0.011 |

HR+ (HR-): hazards ratio for M+ (M-) subgroup, *E*: the number of events at the final analysis,

*p*: observed prevalence of M+ subgroup in the study, *Q*: the empirical probability that the first interaction test is significant.

**Table S1 (continued)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |   |   | *E* = 250 |  | *E* = 350 |
| HR+ | HR- | *p*  | *Q* |   | Standard | Naive correction | *a*1(Zint) | *a*2(Zint) |  | *Q* |   | Standard | Naive correction | *a*1(Zint) | *a*2(Zint) |
| 0.7 | 0.7 | 0.3 | 0.105 | Bias | -0.207 | -0.148 | -0.065 | -0.024 |  | 0.100 | Bias | -0.175 | -0.121 | -0.049 | -0.013 |
|  |  |  |  | Var | 0.007 | 0.011 | 0.018 | 0.027 |  |  | Var | 0.005 | 0.008 | 0.013 | 0.019 |
|  |  |  |  | MSE | 0.05 | 0.033 | 0.022 | 0.028 |  |  | MSE | 0.036 | 0.023 | 0.015 | 0.019 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  | 0.6 | 0.108 | Bias | -0.115 | -0.08 | -0.033 | -0.011 |  | 0.103 | Bias | -0.098 | -0.067 | -0.027 | -0.008 |
|  |  |  |  | Var | 0.007 | 0.009 | 0.012 | 0.015 |  |  | Var | 0.005 | 0.006 | 0.008 | 0.01 |
|  |  |  |  | MSE | 0.021 | 0.015 | 0.013 | 0.015 |  |  | MSE | 0.014 | 0.01 | 0.008 | 0.01 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 0.7 | 0.8 | 0.3 | 0.220 | Bias | -0.162 | -0.102 | -0.018 | 0.020 |  | 0.239 | Bias | -0.136 | -0.083 | -0.012 | 0.020 |
|  |  |  |  | Var | 0.008 | 0.013 | 0.022 | 0.033 |  |  | Var | 0.007 | 0.010 | 0.016 | 0.024 |
|  |  |  |  | MSE | 0.035 | 0.023 | 0.022 | 0.034 |  |  | MSE | 0.025 | 0.017 | 0.017 | 0.025 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  | 0.6 | 0.223 | Bias | -0.086 | -0.052 | -0.006 | 0.012 |  | 0.263 | Bias | -0.075 | -0.046 | -0.009 | 0.005 |
|  |  |  |  | Var | 0.008 | 0.010 | 0.013 | 0.016 |  |  | Var | 0.006 | 0.007 | 0.009 | 0.012 |
|  |  |  |  | MSE | 0.015 | 0.012 | 0.013 | 0.016 |  |  | MSE | 0.011 | 0.009 | 0.009 | 0.012 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 0.7 | 1 | 0.3 | 0.518 | Bias | -0.099 | -0.044 | 0.033 | 0.061 |  | 0.594 | Bias | -0.068 | -0.022 | 0.040 | 0.060 |
|  |  |  |  | Var | 0.013 | 0.020 | 0.035 | 0.050 |  |  | Var | 0.010 | 0.016 | 0.026 | 0.036 |
|  |  |  |  | MSE | 0.023 | 0.022 | 0.036 | 0.054 |  |  | MSE | 0.015 | 0.016 | 0.028 | 0.040 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  | 0.6 | 0.546 | Bias | -0.049 | -0.019 | 0.021 | 0.033 |  | 0.642 | Bias | -0.032 | -0.008 | 0.022 | 0.030 |
|  |  |  |  | Var | 0.009 | 0.012 | 0.016 | 0.020 |  |  | Var | 0.007 | 0.009 | 0.012 | 0.014 |
|  |  |  |  | MSE | 0.012 | 0.012 | 0.017 | 0.021 |  |  | MSE | 0.008 | 0.009 | 0.012 | 0.015 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 0.7 | 1.2 | 0.3 | 0.748 | Bias | -0.052 | -0.003 | 0.063 | 0.080 |  | 0.844 | Bias | -0.027 | 0.008 | 0.055 | 0.063 |
|  |  |  |  | Var | 0.016 | 0.025 | 0.043 | 0.059 |  |  | Var | 0.014 | 0.021 | 0.034 | 0.043 |
|  |  |  |  | MSE | 0.019 | 0.026 | 0.047 | 0.065 |  |  | MSE | 0.015 | 0.021 | 0.037 | 0.047 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  | 0.6 | 0.792 | Bias | -0.020 | 0.003 | 0.034 | 0.040 |  | 0.889 | Bias | -0.009 | 0.008 | 0.029 | 0.031 |
|  |  |  |  | Var | 0.011 | 0.014 | 0.019 | 0.022 |  |  | Var | 0.008 | 0.010 | 0.013 | 0.015 |
|  |  |  |  | MSE | 0.011 | 0.014 | 0.020 | 0.024 |  |  | MSE | 0.008 | 0.010 | 0.014 | 0.016 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

HR+ (HR-): hazards ratio for M+ (M-) subgroup, *E*: the number of events at the final analysis,

*p*: observed prevalence of M+ subgroup in the study, *Q*: the empirical probability that the first interaction test is significant.

# **Table S2. The empirical coverage probability of the 95% confidence interval of hazard ratio in the M+ subgroup: additional scenarios.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|   |   |   | *E* = 250 |  | *E* = 350 |
| HR+ | HR- | *p* | *Q* | Standard | Naive correction | *a*1(Zint) | *a*2(Zint) |  | *Q* | Standard | Naive correction | *a*1(Zint) | *a*2(Zint) |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 0.5 | 0.7 | 0.3 | 0.473 | 0.939 | 0.968 | 0.977 | 0.970 |  | 0.566 | 0.955 | 0.973 | 0.979 | 0.974 |
|  |  | 0.6 | 0.507 | 0.950 | 0.970 | 0.979 | 0.977 |  | 0.601 | 0.958 | 0.972 | 0.979 | 0.975 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 0.5 | 0.8 | 0.3 | 0.665 | 0.964 | 0.974 | 0.977 | 0.974 |  | 0.763 | 0.963 | 0.971 | 0.973 | 0.970 |
|  |  | 0.6 | 0.708 | 0.961 | 0.971 | 0.978 | 0.975 |  | 0.809 | 0.963 | 0.970 | 0.974 | 0.972 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 0.5 | 1 | 0.3 | 0.892 | 0.969 | 0.973 | 0.974 | 0.974 |  | 0.948 | 0.963 | 0.970 | 0.973 | 0.973 |
|  |  | 0.6 | 0.922 | 0.962 | 0.966 | 0.970 | 0.970 |  | 0.967 | 0.953 | 0.956 | 0.959 | 0.959 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 0.5 | 1.2 | 0.3 | 0.968 | 0.964 | 0.970 | 0.973 | 0.973 |  | 0.992 | 0.954 | 0.961 | 0.967 | 0.967 |
|  |  | 0.6 | 0.981 | 0.951 | 0.955 | 0.958 | 0.957 |  | 0.996 | 0.953 | 0.956 | 0.957 | 0.957 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 0.6 | 0.7 | 0.3 | 0.247 | 0.890 | 0.958 | 0.975 | 0.967 |  | 0.263 | 0.902 | 0.965 | 0.977 | 0.973 |
|  |  | 0.6 | 0.259 | 0.919 | 0.967 | 0.981 | 0.979 |  | 0.283 | 0.916 | 0.967 | 0.980 | 0.977 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 0.6 | 0.8 | 0.3 | 0.404 | 0.940 | 0.970 | 0.978 | 0.973 |  | 0.494 | 0.945 | 0.968 | 0.976 | 0.970 |
|  |  | 0.6 | 0.428 | 0.944 | 0.974 | 0.983 | 0.979 |  | 0.517 | 0.948 | 0.971 | 0.981 | 0.977 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 0.6 | 1 | 0.3 | 0.710 | 0.964 | 0.972 | 0.975 | 0.971 |  | 0.812 | 0.965 | 0.970 | 0.972 | 0.969 |
|  |  | 0.6 | 0.755 | 0.958 | 0.968 | 0.973 | 0.970 |  | 0.858 | 0.963 | 0.969 | 0.974 | 0.972 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 0.6 | 1.2 | 0.3 | 0.886 | 0.969 | 0.973 | 0.974 | 0.973 |  | 0.972 | 0.966 | 0.973 | 0.975 | 0.975 |
|  |  | 0.6 | 0.918 | 0.958 | 0.962 | 0.965 | 0.965 |  | 0.972 | 0.956 | 0.959 | 0.962 | 0.962 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 0.7 | 0.7 | 0.3 | 0.105 | 0.779 | 0.949 | 0.976 | 0.968 |  | 0.100 | 0.769 | 0.957 | 0.983 | 0.970 |
|  |  | 0.6 | 0.108 | 0.848 | 0.938 | 0.974 | 0.969 |  | 0.103 | 0.855 | 0.957 | 0.984 | 0.979 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 0.7 | 0.8 | 0.3 | 0.220 | 0.940 | 0.970 | 0.978 | 0.973 |  | 0.239 | 0.945 | 0.968 | 0.976 | 0.970 |
|  |  | 0.6 | 0.223 | 0.944 | 0.974 | 0.983 | 0.979 |  | 0.263 | 0.948 | 0.971 | 0.981 | 0.977 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 0.7 | 1 | 0.3 | 0.518 | 0.944 | 0.965 | 0.972 | 0.967 |  | 0.594 | 0.957 | 0.971 | 0.976 | 0.971 |
|  |  | 0.6 | 0.546 | 0.952 | 0.973 | 0.980 | 0.976 |  | 0.642 | 0.960 | 0.975 | 0.979 | 0.977 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 0.7 | 1.2 | 0.3 | 0.748 | 0.963 | 0.970 | 0.973 | 0.969 |  | 0.844 | 0.967 | 0.972 | 0.974 | 0.972 |
|  |  | 0.6 | 0.792 | 0.959 | 0.969 | 0.973 | 0.973 |  | 0.889 | 0.962 | 0.967 | 0.970 | 0.969 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |

HR+ (HR-): hazards ratio for M+ (M-) subgroup, *E*: the number of events at the final analysis,

*p*: observed prevalence of M+ subgroup in the study, *Q*: the probability that the first preliminary interaction test is significant.