

Sensitivity analysis

We chose to show a simple and easily interpretable model in the manuscript, with all continuous variables categorised into factor variables. In a sensitivity analysis we performed a logistic regression without this categorisation. This model contained the same variables as the model in the manuscript (PCG cancer, PCG cardiac disease, PCG pain, emergency index admission, number of emergency visits, costs specialists, costs hospital outpatient, costs laboratory, costs therapeutic devices, costs physiotherapy, number of outpatient visits, sex, age group, geographical region). All continuous variables (number of emergency visits, costs specialists, costs hospital outpatient, costs laboratory, costs therapeutic devices, costs physiotherapy, and number of outpatient visits) were transformed (Yeo-Johnson transformation) and expanded in order to fit restricted cubic splines (with three knots). Age was not available as a continuous variable due to anonymisation measures.

The estimated coefficients of this model are in Table 7. Its discrimination was slightly better than the one in the model presented in the manuscript. It achieved an AUC of 0.61 (95%-CI: 0.60 - 0.62). This model was inferior to the model in the manuscript in terms of the calibration. Calibration-in-the-large was 1.83 (95%-CI: -0.15 - 3.82), the calibration slope was 0.78 (95%-CI: 0.56 - 0.99).

Table 7 Estimated coefficients logistic regression model with continuous variables

Term	OR	95% CI	P
(Intercept)	0.04	(0.03 - 0.04)	<0.001
PCG cancer (= yes)	1.13	(1.00 - 1.27)	0.058
PCG cardiac disease (= yes)	1.07	(1.02 - 1.13)	0.005
PCG pain (= yes)	1.15	(1.08 - 1.22)	<0.001
emergency index admission (= yes)	1.49	(1.43 - 1.56)	<0.001
rcs(number of emergency visits)	1.05	(0.99 - 1.12)	0.109
rcs(number of emergency visits)'	0.99	(0.95 - 1.04)	0.805
rcs(costs specialists)	0.94	(0.90 - 0.98)	0.007
rcs(costs specialists)'	1.01	(0.96 - 1.07)	0.672
rcs(costs hospital outpatient)	0.88	(0.83 - 0.93)	<0.001
rcs(costs hospital outpatient)'	1.25	(1.17 - 1.34)	<0.001
rcs(costs laboratory)	1.02	(0.97 - 1.07)	0.412
rcs(costs laboratory)'	1.11	(1.06 - 1.15)	<0.001
rcs(costs therapeutic devices)	1.01	(0.95 - 1.07)	0.834
rcs(costs therapeutic devices)'	1.06	(0.99 - 1.14)	0.080
rcs(costs physiotherapy)	1.00	(0.86 - 1.18)	0.965

rsc(costs physiotherapy)'	0.78	(0.39 - 1.59)	0.502
rsc(number of outpatient visits)	0.99	(0.98 - 1.00)	0.021
rsc(number of outpatient visits)'	1.01	(1.00 - 1.02)	0.021
sex (= male)	1.29	(1.24 - 1.35)	<0.001
age group (= 30 - 39)	0.98	(0.87 - 1.11)	0.802
age group (= 40 - 49)	1.17	(1.05 - 1.31)	0.006
age group (= 50 - 59)	1.26	(1.13 - 1.40)	<0.001
age group (= 60 - 69)	1.42	(1.28 - 1.58)	<0.001
age group (= 70 - 79)	1.40	(1.26 - 1.56)	<0.001
age group (= 80 - 89)	1.54	(1.37 - 1.72)	<0.001
age group (= 90 - ..)	1.29	(1.09 - 1.54)	0.004
region (= Lake Geneva)	0.74	(0.68 - 0.79)	<0.001
region (= Midland)	0.92	(0.87 - 0.98)	0.010
region (= Northwest)	0.94	(0.88 - 1.01)	0.082
region (= East)	1.05	(0.98 - 1.12)	0.145
region (= Ticino)	0.95	(0.88 - 1.02)	0.178
region (= Central)	0.81	(0.74 - 0.90)	<0.001