Developing a Prediction Model for 7-year and 10-year All-cause Mortality Risk in Type 2 Diabetes Using a Hospital-based Prospective Cohort Study

Sherry Yueh-Hsia Chiu\textsuperscript{1,2}, Ying Isabel Chen\textsuperscript{3}, Jui-fen Rachel Lu\textsuperscript{4,5}, Soh-Ching Ng\textsuperscript{6}, Chih-Hung Chen\textsuperscript{6}

1. Department of Health Care Management, College of Management; and Healthy Aging Research Center, Chang Gung University, Taiwan
2. Division of Hepato-gastroenterology, Department of Internal Medicine, Kaohsiung Chang Gung Memorial Hospital, Taiwan
3. Graduate Institute of Epidemiology and Preventive Medicine, College of Public Health, National Taiwan University, Taoyuan, Taiwan
4. Graduate Institute of Business and Management and Department of Health Care Management, College of Management, Taoyuan, Chang Gung University
5. Department of Radiation Oncology, Linkou Chang Gung Memorial Hospital, Linkou, Taiwan
6. Division of Endocrinology and Metabolism, Department of Internal Medicine, Keelung Chang Gung Memorial Hospital, Keelung; Chang Gung University, Taiwan
Figure S1. Proposed study scenario and application
2007-2013
All outpatients aged ≥18y/o with one diabetes mellitus diagnosis code
34164 subjects

Excluded type 1 diabetes mellitus
271 subjects

Outpatients with type 2 diabetes mellitus code
33893 subjects

Excluded individuals had <3 clinic visits with type 2 diabetes mellitus diagnosis/year
15680 subjects

Type 2 diabetes mellitus
(criterion ≥ 3 clinic visits with Type 2 diabetes mellitus diagnosed code/year)
18213 subjects

Excluded subjects with follow-up <1 year
11 subjects

Individuals with type 2 diabetes mellitus
18202 subjects

Linking with National Mortality Registry of 2013
(Total 2779 deaths)

Linking with National Mortality Registry of 2016
(Total 4561 deaths)

Figure S2. Flow diagram for study subjects
**Figure S3(A).** Proportional hazards assumption checking for 7-year follow-up

**Figure S3(B).** Proportional hazards assumption checking for 10-year follow-up
18,202 subjects

Random sampling

50% subjects: Training data
9101(T)

Multivariable Cox regression
Coefficients (9101(T))

Predictive validity
1. Time-dependent ROC
2. AUC
3. Cumulative mortality

50% subjects: Validation data
9101(V)

Multivariable Cox regression
Coefficients (9101(V))

From observed data
1. Time-dependent ROC
2. AUC
3. Cumulative mortality

9101(T): training dataset
9101(V): validation dataset

Figure S4. Flowchart for model cross-validation
Figure S5(A). Cumulative all-cause mortality for 7-year follow-up based on 9101 validation subjects

Figure S5(B). Cumulative all-cause mortality for 10-year follow-up based on 9101 validation subjects
Figure S6(A). ROC and AUC for cross-validation at the 2-year time point based on the 7-year follow-up

From observed data, AUC= 0.8077
Predictive validity, AUC= 0.8044

Figure S6(B). ROC and AUC for cross-validation at the 4-year time point based on the 7-year follow-up

From observed data, AUC= 0.7905
Predictive validity, AUC= 0.7925

Figure S6(C). ROC and AUC for cross-validation at the 6-year time point based on the 7-year follow-up

From observed data, AUC= 0.7906
Predictive validity, AUC= 0.7907
**Figure S7(A).** ROC and AUC for cross-validation at the 2-year time point based on the 10-year follow-up

**Figure S7(B).** ROC and AUC for cross-validation at the 4-year time point based on the 10-year follow-up

**Figure S7(C).** ROC and AUC for cross-validation at the 6-year time point based on the 10-year follow-up

**Figure S7(D).** ROC and AUC for cross-validation at the 8-year time point based on the 10-year follow-up