Supplementary Materials

Translation of Risk Estimates into a Risk Calculator

To transform our model coefficients (examples provided below) into an individualized 3-year mortality estimate the equation below was used:\footnote{1}:

\[
\hat{p}_{ez} = \frac{\exp[\hat{a} + \hat{\beta}_1 \times e + \hat{\beta}_2 \times z]}{1 + \exp[\hat{a} + \hat{\beta}_1 \times e + \hat{\beta}_2 \times z]}
\]

3-year mortality risk for a dialysis patient is derived from the formula below,

\[-2.9578(\text{Baseline risk}) + 0.0067(\text{Female}) + 0.0388(\text{Age}) -0.2990(\text{Black race}) -0.6111(\text{Other Race}) + 0.4737(\text{Cardiovascular disease}) -0.4696(\text{Hypertension}) + 0.0169(\text{Diabetes})\]

3-year mortality among transplant patients is derived from the formula below,

\[-5.4292(\text{Baseline risk}) -0.0475(\text{Female}) + 0.0382(\text{Age}) -0.0261(\text{Black race}) -0.508 (\text{Other race}) + 0.3369(\text{Cardiovascular disease}) -0.2(\text{Hypertension}) + 0.4013 (\text{Diabetes}) + 0.136 (6-12 months on dialysis) + 0.4906 (>12 months on dialysis)\]

3-year mortality among deceased donor transplantation is derived from the formula below:

\[-5.0828(\text{Baseline risk}) -0.1517(\text{Female}) + 0.0391(\text{Age}) + 0.0258(\text{Black race}) -0.5563 (\text{Other race}) + 0.3080(\text{Cardiovascular disease}) -0.1801(\text{Hypertension}) + 0.2730 (\text{Diabetes}) + 0.1732(6-12 months on dialysis) + 0.1605 (>12 months on dialysis)\]

3-year mortality among living donor transplantation is derived from the formula below:

\[-5.6803(\text{Baseline risk}) -0.1231(\text{Female}) + 0.0364(\text{Age}) -0.1563(\text{Black race}) -0.4707 (\text{Other race}) + 0.5908(\text{Cardiovascular disease}) -0.1417(\text{Hypertension}) + 0.4418(\text{Diabetes}) + 0.0614(6-12 months on dialysis) + 0.3202 (>12 months on dialysis)\]

\textbf{Note:} Risk estimates for living versus deceased donor kidney transplant assumes you would receive a kidney from either donor type on the same day.

\textit{Where baseline risk=1; male=0, female= 1; 1=yes and 0=no for Black race, other race, cardiovascular disease, hypertension, diabetes, 6-12 months on dialysis, and >12 months on dialysis. Age is modeled as a continuous integer variable.}

\textbf{Reference}

1. Muller CJ and MacLehose RF. Estimating predicted probabilities from logistic regression: different methods correspond to different target populations. International Journal of Epidemiology. 2014; 43 (3); 962–970.
**Supplementary Table 1:** Checklist of recommendations for reporting of observational studies using the REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement

| Item No | STROBE items | RECORD items | Reported |
|---------|--------------|--------------|----------|
| **Title and abstract** | (a) Indicate the study's design with a commonly used term in the title or the abstract. (b) Provide in the abstract an informative and balanced summary of what was done and what was found. | (1.1) The type of data used should be specified in the title or abstract. When possible, the name of the databases used should be included. (1.2) If applicable, the geographic region and time frame within which the study took place should be reported in the title or abstract. (1.3) If linkage between databases was conducted for the study, this should be clearly stated in the title or abstract. | Abstract |
| **Introduction** | | | |
| Background/ rationale | 2 | Explain the scientific background and rationale for the investigation being reported. | Introduction |
| Objectives | 3 | State specific objectives, including any prespecified hypotheses. | Introduction |
| **Methods** | | | |
| Study design | 4 | Present key elements of study design early in the paper. | Methods: Design and Setting |
| Setting | 5 | Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection. | Methods: Study Population, Data Sources, & Statistical Analysis |
| Participants | 6 | (a) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up. (b) For matched studies, give matching criteria and number of exposed and unexposed. | (6.1) The methods of study population selection (such as codes or algorithms used to identify subjects) should be listed in detail. If this is not possible, an explanation should be provided. (6.2) Any validation studies of the codes or algorithms used to select the population should be referenced. If validation was conducted for this study and not published elsewhere, detailed methods and results should be provided. (6.3) If the study involved linkage of databases, consider use of a flow diagram or other graphical display to demonstrate the data linkage process, including the number of individuals with linked data at each stage. | Methods: Study Populations, Data Sources, & Supplementary Table 3 |
| Variables | 7 | Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable. | (7.1) A complete list of codes and algorithms used to classify exposures, outcomes, confounders, and effect modifiers should be provided. If these cannot be reported, an explanation should be provided. | Methods: Study Variables, Supplementary Table 3 |
| Data sources/ measurement | 8 | For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group. |
|---------------------------|---|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Bias                      | 9 | Describe any efforts to address potential sources of bias. |
| Study size                | 10| Explain how the study size was arrived at. |
| Quantitative variables    | 11| Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why. |
| Statistical methods       | 12| (a) Describe all statistical methods, including those used to control for confounding.  
(b) Describe any methods used to examine subgroups and interactions.  
(c) Explain how missing data were addressed.  
(d) If applicable, explain how loss to follow-up was addressed.  
(e) Describe any sensitivity analyses. |
| Data access and cleaning methods | N/A | (12.1) Authors should describe the extent to which the investigators had access to the database population used to create the study population.  
(12.2) Authors should provide information on the data cleaning methods used in the study. |
| Linkage                   | N/A | (12.3) State whether the study included person-level, institutional-level, or other data linkage across two or more databases. The methods of linkage and methods of linkage quality evaluation should be provided. |

**Results**

| Participants              | 13| (a) Report numbers of individuals at each stage of study—e.g. numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analyzed.  
(b) Give reasons for non-participation at each stage.  
(c) Consider use of a flow diagram. |
|----------------------------|---|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Descriptive data           | 14| (a) Give characteristics of study participants (e.g. demographic, clinical, social) and information on exposures and potential confounders.  
(b) Indicate number of participants with missing data for each variable of interest.  
(c) Summarize follow-up time (e.g. average and total amount). |
| Outcome data | 15 | Report numbers of outcome events or summary measures over time. |
|-------------|----|---------------------------------------------------------------|
| **Main results** | 16 | (a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g. 95% confidence interval). Make clear which confounders were adjusted for and why they were included. (b) Report category boundaries when continuous variables were categorized. (c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period. |
| **Other analyses** | 17 | Report other analyses done (e.g. analyses of subgroups and interactions, and sensitivity analyses). |
| **Key results** | 18 | Summarize key results with reference to study objectives. |
| **Limitations** | 19 | Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias. |
| **Interpretation** | 20 | Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence. |
| **Generalizability** | 21 | Discuss the generalizability (external validity) of the study results. |
| **Other information** | | |
| **Funding** | 22 | Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based. |
| **Accessibility of protocol, raw data, and programming code** | N/A | (22.1) Authors should provide information on how to access any supplemental information such as the study protocol, raw data, or programming code. |

Reference: Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. PLoS medicine. 2015;12(10):e1001885.
Supplementary Table 2: Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement for prediction model validation

| Section/Topic | Checklist Item | Page |
|---------------|----------------|------|
| **Title and abstract** | | |
| Title | Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted. | Title |
| Abstract | Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions. | Abstract |
| **Introduction** | | |
| Background and objectives | Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models. | Introduction |
| | Specify the objectives, including whether the study describes the development or validation of the model or both. | Introduction |
| **Methods** | | |
| Source of data | Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable. | Design and Setting |
| | Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up. | Study Populations, Study Variables |
| Participants | Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres. | Design and Setting |
| | Describe eligibility criteria for participants. | Study Populations |
| | Give details of treatments received, if relevant. | NA |
| Outcome | Clearly define the outcome that is predicted by the prediction model, including how and when assessed. | Study Variables |
| | Report any actions to blind assessment of the outcome to be predicted. | NA |
| Predictors | Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured. | Data Sources, Study Variables |
| | Report any actions to blind assessment of predictors for the outcome and other predictors. | NA |
| Sample size | Explain how the study size was arrived at. | Supplementary Figures 1 and 2 |
| Missing data | Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method. | Study Variables |
| | For validation, describe how the predictions were calculated. | Study Variables, Statistical Analysis |
| Statistical analysis methods | 10d | Specify all measures used to assess model performance and, if relevant, to compare multiple models. | Statistical Analysis |
|-------------------------------|-----|------------------------------------------------------------------------------------------------|---------------------|
|                               | 10e | Describe any model updating (e.g., recalibration) arising from the validation, if done.          | Statistical Analysis |
| Risk groups                  | 11  | Provide details on how risk groups were created, if done.                                        | NA                  |
| Development vs. validation    | 12  | For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors. | Baseline Characteristics, Discussion |

### Results

#### Participants

| 13a | Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful. | Supplementary Figures 1 and 2, Baseline Characteristics |
|-----|------------------------------------------------------------------------------------------------|--------------------------------------------------|
| 13b | Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome. | Study Variables, Baseline Characteristics, Table 1 |
| 13c | For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome). | Baseline Characteristics |

#### Model performance

| 16  | Report performance measures (with CIs) for the prediction model. | Predictive Model Discrimination, Predictive Model Calibration, Table 2, Figures 1-4, Supplementary Figures 3-5 |

#### Model-updating

| 17  | If done, report the results from any model updating (i.e., model specification, model performance). | Predictive Model Calibration, Figures 1 and 2, Supplementary Figures 4 and 5 |

### Discussion

#### Limitations

| 18  | Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data). | Discussion |

#### Interpretation

| 19a | For validation, discuss the results with reference to performance in the development data, and any other validation data. | Discussion |
| 19b | Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence. | Discussion |

#### Implications

| 20  | Discuss the potential clinical use of the model and implications for future research. | Discussion |
| Other information | 21 | F22 |
|------------------|----|-----|
| Supplementary information | Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets. | Availability of data and material |
| Funding | Give the source of funding and the role of the funders for the present study. | Funding |

Reference: Moons KG, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med*. 2015;162(1):W1-73.
Supplementary Table 3. Coding definitions for demographic and comorbid conditions

| Characteristic                                | Database | Codes                                                                 |
|-----------------------------------------------|----------|----------------------------------------------------------------------|
|                                | **Inclusion Criteria**                                                                                   |
| Chronic Dialysis                            | CORR     | Treatment_Code: 111, 112, 113, 121, 122, 123, 131, 132, 133, 211, 221, 231, 311, 312, 313, 321, 322, 323, 331, 332, 333, 413, 423, 433, 443, 453, 141, 151, 152, 241, 242, 251, 252, 443, 453 |
| Transplant Recipient                        | CORR     | Treatment_Code: 171                                                 |
|                                              |          | Transplanted_Organ_Type_Code: 10, 11, 12, 18, 19                    |
|                                | **Exclusion Criteria**                                                                                   |
| Previous Transplant                          | CORR     | Treatment_Code: 171, 181                                             |
|                                              |          | Graft_Num: >=1                                                       |
| Prior Dialysis                               | CORR     | Treatment_Code: 111, 112, 113, 121, 122, 123, 131, 132, 133, 211, 221, 231, 311, 312, 313, 321, 322, 323, 331, 332, 333, 413, 423, 433, 443, 453, 141, 151, 152, 241, 242, 251, 252, 443, 453 |
| Simultaneous Multi-Organ Transplant          | CORR     | Transplanted_Organ_Type_Code[1-3] is not missing and equal to a code other than 10, 11, 12, 18, 19 |
|                                | **Demographics**                                                                                          |
| Age                                           | RPDB     |                                                                      |
| Sex                                           | RPDB     |                                                                      |
| Race                                          | CORR     | Racial_Origin_Code:                                                  |
|                                              |          | Caucasian: 01                                                        |
|                                              | ORRS     | Black: 03                                                            |
|                                              |          | Other: 99, 02, 05, 08, 09, 10, 11                                     |
|                                              |          | Unknown: 98                                                          |
|                                              |          | RaceCD:                                                              |
|                                              | ORRS     | Caucasian: 01                                                        |
|                                              |          | Black: 12                                                            |
|                                              | ORRS     | Other: 99, 02, 05, 08, 09, 10, 11                                     |
|                                              | ORRS     | Unknown: 98                                                          |
| Rural                                         | RPDB     |                                                                      |
| Income                                        | RPDB     |                                                                      |
| Patient Clinical Characteristics          | CORR |                          |
|------------------------------------------|------|--------------------------|
| **Cause of end-stage kidney disease**    |      | Primary_Diagnosis_Kidney: |
|                                          |      | Glomerulonephritis/Autoimmune: 05, 06, 07, 08, 09, 10, 12, 13, 14, 15, 16, 19, 73, 74, 84, 85, 86, 88 |
|                                          |      | Cystic Kidney Disease: 40, 41, 42, 43, 49 |
|                                          |      | Diabetes: 80, 81 |
|                                          |      | Renal Vascular Disease: 70, 71, 72, 79 |
|                                          |      | Other: 20, 21, 22, 23, 24, 25, 29, 30, 31, 32, 33, 39, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 66, 78, 82, 83, 87, 89, 90, 91, 92, 93, 94, 95, 96, 97, 99 |
|                                          |      | Unknown: 00, 98 |
| **Dialysis Vintage**                     | CORR | Treatment_Code:          |
|                                          |      | Transplant: 171          |
|                                          |      | Dialysis: 060, 111, 112, 113, 121, 122, 123, 131, 132, 133, 141, 151, 152, 211, 221, 231, 241, 242, 251, 252, 311, 312, 313, 321, 322, 323, 331, 332, 333, 413, 423, 433, 443, 453 |
| **Dialysis Modality**                    | CORR | Treatment_Code:          |
|                                          |      | Hemodialysis: 060, 111, 112, 113, 121, 122, 123, 131, 132, 133, 211, 221, 231, 241, 242, 251, 252, 311, 312, 313, 321, 322, 323, 331, 332, 333, 413, 423, 433 |
|                                          |      | Peritoneal: 141, 151, 152, 241, 242, 251, 252, 443, 453 |
| **Donor Type**                           | CORR | Donor_Type_Code:         |
|                                          |      | Living: 02, 03, 04, 05, 06, 07, 10, 11, 15 |
|                                          |      | Deceased: 01 |
|                                          |      | Missing/Unknown: 98 |
| **Body Mass Index**                      | CORR | Initial_Height           |
|                                          |      | Initial_Weight           |
|                                          |      | Calculate BMI:           |
|                                          |      | Height2=initial_height*1/100 |
|                                          |      | Weight2=initial_weight*1  |
|                                          |      | BMI=weight2/(height2**2)  |
| **Cardiovascular Disease**               | CIHI/| ICD-9: 425, 5184, 514, 428, 412, 410, 413, 414, 4292, 4296, 4297, 411, 430, 431, 432, 434, 435, 436, 3623, 4402, 4408, 4409, 5571, 4439, 444, 410, 491, 492, 496, 4273 |
|                                          | NACRS| ICD-10: I500, I501, I509, I255, J81, I20, I21, I22, I23, I24, I25, Z955, Z958, Z959, R931, T822, I62, I630, I631, I632, I633, I634, I635, I638, I639, I64, H341, I600, |
| Group | Abbreviation | Code Details |
|-------|--------------|-------------|
| OHIP | I601, I602, I603, I604, I605, I606, I607, I609, I61, G450, G451, G452, G453, G458, G459, H340, I700, I702, I708, I709, I731, I738, I739, K551, I21, I22, J41, J43, J44, J48, CCP: 4961, 4962, 4963, 4964, 4801, 4802, 4803, 4804, 4805, 481, 482, 483, 5125, 5129, 5014, 5016, 5018, 5028, 5038, 5126, 5159, CCI: 1HP53, 1HP55, 1HZ53GRFR, 1HZ53LAFR, 1HZ53SYFR, 1IJ26, 1IJ27, 1IJ54, 1IJ57, 1IJ50, 1IJ76, 1KA76, 1KA50, 1KE76, 1KG50, 1KG57, 1KG76MI, 1KG87, 1IA87LA, 1IB87LA, 1IC87LA, 1ID87, 1KA87LA, 1KE57, OHIP Fee Code: R701, R702, Z429, R741, R742, R743, G298, E646, E651, E652, E654, E655, G262, Z434, Z448, R787, R780, R797, R804, R809, R809, R815, R936, R783, R784, R785, E626, R814, R786, R937, R860, R861, R855, R856, R933, R934, R791, R794, R813, R867, E649, OHIP Diagnosis Code: 428, 410, 412, 413 |
| Hypertension | CIHI | ICD-9: 401, 402, 403, 404, 405 |
| OHIP | OHIP | OHIP Diagnosis Code: 401, 402, 403 |
| Diabetes | CIHI | ICD-9: 250 |
| OHIP | OHIP | OHIP Diagnosis Code: 250 |
| Charlson Comorbidity Index | CIHI | OHIP Fee Code: K045, K046, K029, K030, Q040 |

### Outcome

| Mortality | RPDB |
|-----------|------|

**Abbreviations:** CCI, Canadian Classification of Interventions; CCP, Canadian Classification of Diagnostic, Therapeutic and Surgical Procedures; CIHI, Canadian Institute for Health Information; CORR, Canadian Organ Replacement Registry; ICD, International Classification of Disease; NACRS, National Ambulatory Care Reporting System; OHIP, Ontario Health Insurance Plan; ORRS, Ontario Renal Reporting System; RPDB, Registered Persons Database
**Supplementary Table 4.** C-statistics for mortality prediction for maintenance dialysis patients and kidney transplant recipients restricting cohort entry years from 2004 to 2013

| Models                  | c-statistic (95% CI) |
|-------------------------|----------------------|
| Maintenance Dialysis    | 0.70 (0.69, 0.71)    |
| Kidney Transplant Recipients | 0.72 (0.68, 0.75) |
| Deceased Donor Transplant | 0.68 (0.64, 0.72) |
| Living Donor Transplant  | 0.71 (0.65, 0.77)    |

Abbreviation: CI, confidence interval
Supplementary Figure 1. Maintenance dialysis cohort

30,795 Ontario maintenance dialysis patients

4618 individuals excluded during data cleaning (i.e., missing health card number, missing sex, missing date of birth, non-Ontario resident, death before index date, aged <18 or >80 years)

26,177 individuals after data cleaning

Excluded (n=3657):
1694 with evidence of previous organ transplant, including kidney
1963 with evidence of prior chronic dialysis

22,520 individuals included in maintenance dialysis cohort
Supplementary Figure 2. Kidney transplant recipient cohort

5502 Ontario kidney transplant recipients

189 individuals excluded during data cleaning (i.e., missing health card number, missing sex, missing date of birth, non-Ontario resident, death before index date, aged <18 or >80 years)

5313 individuals after data cleaning

Excluded (n=808):
537 with evidence of previous organ transplant, including kidney
271 with evidence of simultaneous multi-organ transplant

4505 individuals included in kidney transplant cohort
Supplementary Figure 3. Calibration plot comparing the observed and predicted mortality probability for recipients of a living donor kidney transplant.
**Supplementary Figure 4.** Observed and predicted mortality risk in recipients of a deceased donor kidney after intercept recalibration. *In accordance with ICES privacy policies, cell sizes less than or equal to five cannot be reported.*
**Supplementary Figure 5.** Calibration plot comparing the observed and predicted mortality probability for recipients of a deceased donor kidney before (a) and after (b) intercept recalibration.