**Online Supplement**

**Prediction of vitamin D deficiency in older adults:**

**the role of machine learning models**

John D. Sluyter1, Yoshihiko Raita2, Kohei Hasegawa2, Ian R. Reid3, Robert Scragg1, Carlos A. Camargo Jr2

*1School of Population Health, University of Auckland, Auckland, New Zealand; 2Department of Emergency Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, USA; 3Department of Medicine, Faculty of Medical and Health Sciences, University of Auckland, Auckland, New Zealand*

**Abbreviations used in the Online Supplement:**

25(OH)D = deseasonalized 25-hydroxyvitamin D.

**Contents**

**Supplementary Table 1. 3**

Questions for potential predictors derived from questionnaire data

**Supplementary Figure 1.**  **4**

Prediction performance of simple models for detection, in the test set, of low serum 25(OH)D (deseasonalized 25-hydroxyvitamin D): a) <50 nmol/L (ROC curves), b) <50 nmol/L (decision curves), c) <25 nmol/L (ROC curves), d) <25 nmol/L (decision curves). For decision curves, the net benefit associated with not testing anyone for vitamin D deficiency and testing all are given by the black horizontal lines (net benefit=0) and grey angled lines, respectively

**Supplementary Figure 2. 5**

Mean serum 25(OH)D, in the test set, of low 25(OH)D: a) <50 nmol/L with simple models, b) <50 nmol/L with augmented models, c) <25 nmol/L with simple models, d) <25 nmol/L with augmented models

**Supplementary Figure 3. 6**

Variable importance of predictors in the simple models for predicting low serum 25(OH)D (deseasonalized 25-hydroxyvitamin D): a) <50 nmol/L and, b) <25 nmol/L. The variable importance is a scaled measure to have a maximum value of 100

**Supplementary Table 1.** Questions for potential predictors derived from questionnaire data

|  |  |  |
| --- | --- | --- |
| Predictor | Question from questionnaire | Question responses (n, type) |
| Age (years) | Today’s date (not in questionnaire; recorded by interviewers) | OE (date) |
| What is your date of birth? | OE (date) |
| Female sex | You are male / female ...? | 2 MC |
| Ethnicity | Which ethnic group do you most identify with? | 32 MC |
| Home setting | What is the setting of your home? | 5 MC |
| Marital status | What is your current marital status? | 4 MC |
| Employment status | Are you in paid work now? | 2 MC |
| Self-rated general health | Looking at this showcard, in general, would you say your health is: | 5 LS |
| Diabetes mellitus | Have you ever been told by a doctor that you have diabetes?\* | 2 MC |
| Cardiovascular disease | Have you ever been told by a doctor that you have had a heart attack? | 2 MC |
| Have you ever been told by a doctor that you have had angina? | 2 MC |
| Have you ever been told by a doctor that you have heart failure? | 2 MC |
| Have you ever been told by a doctor that you have had a “mini-stroke” or transient ischaemic attack? | 2 MC |
| Have you ever been told by a doctor that you have had a stroke? | 2 MC |
| Have you ever been told by a doctor that you have carotid artery stenosis (blocked neck arteries) or had surgery on your carotid (neck) arteries? | 2 MC |
| Hyperlipidemia | Have you ever been told by a doctor that you have high cholesterol levels in your blood? | 2 MC |
| Urgent treatment for asthma or chronic bronchitis/emphysema in past 12 months | How many times did you go to a doctor’s office, clinic or surgery over the past 12 months for urgent treatment of your asthma or chronic bronchitis/emphysema? | OE (n) |
| How many times did you go to an after-hours medical centre, Hospital A&E or Emergency Department over the past 12 months for urgent treatment of your asthma or chronic bronchitis/emphysema? | OE (n) |
| How many times were you admitted to a hospital over the past 12 months for urgent treatment of your asthma or chronic bronchitis/emphysema? | OE (hrs) |
| Antihypertensive medication | Are you currently taking pills regularly for high blood pressure? | 2 MC |
| Taking vitamin D supplements | Do you currently take any vitamin D or calcium supplements? | 2 MC |
| Please can you show me all your vitamins and supplements? | OE |
| Sun exposure (hrs/day) | In the last 3 months, how many hours each day, did you usually spend outdoors in the sun (including work): |  |
|  | On an average week day? | OE (hrs) |
|  | On an average weekend day? | OE (hrs) |
| Total physical activity (hrs/week) | In a typical week during the past 3 months, how many hours did you spend on each of the following activities? | OE (hrs) |
|  | Walking, including walking to work, shopping and leisure | OE (hrs) |
|  | Cycling, including cycling to work and during leisure time | OE (hrs) |
|  | Gardening | OE (hrs) |
|  | Housework such as cleaning, washing cooking, child care | OE (hrs) |
|  | Do it yourself (this includes home maintenance, …) | OE (hrs) |
|  | Other physical exercise such as keeping fit, aerobics, swimming, jogging | OE (hrs) |
| Vigorous physical activity (hrs/week) | In a typical week during the past 3 months did you practise any of these activities vigorously enough to cause sweating or faster heartbeat? | 2 MC |
| For how many hours per week in total did you do such vigorous physical activity? | OE (hrs) |
| Current-smoker (%) | Which of these best describes you? | 3 MC |
| No alcohol intake in past 12 months (%) | How often have you had a drink containing alcohol in the last 12 months? | 7 LS |

hrs = hours; LS = Likert-scale (ordinal responses); MC = multiple-choice; n = number; OE = open-ended. \*For females, excludes diabetes during pregnancy.

**Supplementary Figure 1: Simple models**

**a) <50 nmol/L, ROC curves b) <50 nmol/L, decision curves**

 

**c) <25 nmol/L, ROC curves d) <25 nmol/L, decision curves**

 



**Supplementary Figure 2**

**a) <25 nmol/L, simple models b) <25 nmol/L, augmented models**

 

**c) <50 nmol/L, simple models d) <50 nmol/L, augmented models**

 

**Supplementary Figure 3: Simple models**

**a) <50 nmol/L**

**b) <25 nmol/L**