**Caries Risk Prediction models in medical healthcare setting**

Tosha Ashish Kalhan1, Carolina Un Lam2,Bindu Karunakaran1, Pui Ling Chay3, Chai Kiat Chng3, Rahul Nair4,16, Yung Seng Lee2,5,6, Mary Chong Foong Fong2,7, Yap Seng Chong2,8, Kenneth Kwek3, Seang-Mei Saw7, Lynette Shek2,5,9, Fabian Yap10,11,12, Kok Hian Tan3,13, Keith M Godfrey14, Jonathan Huang15, Chin-Ying Stephen Hsu1#

**Author affiliations**

1Faculty of Dentistry, National University of Singapore, Singapore

2Chief Dental Officer’s Office, Ministry of Health, College of Medicine Building, Singapore

3Dental Service, KK Women’s and Children’s Hospital, Singapore

4Radboud University Medical Centre, Radboud Institute for Health Sciences, Department of Dentistry - Quality and Safety of Oral Healthcare, Nijmegen, the Netherlands

5Department of Paediatrics, Yong Loo Lin School of Medicine, National University of Singapore, Singapore

6Division of Paediatric Endocrinology and Diabetes, Khoo Teck Puat-National University Children’s Medical Institute, National University Hospital, National University Health System, Singapore

7Saw Swee Hock School of Public Health, National University of Singapore, Singapore

8Department of Obstetrics & Gynaecology, Yong Loo Lin School of Medicine, National University of Singapore

9Division of Paediatric Allergy, Immunology & Rheumatology, National University Hospital, Singapore

10Department of Paediatrics, KK Women’s and Children’s Hospital, Singapore

11Department of Paediatrics, Duke-National University of Singapore Graduate Medical School, Singapore

12Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore

13Duke-National University of Singapore Graduate Medical School, Singapore

14MRC Lifecourse Epidemiology Unit & NIHR Southampton Biomedical Research Centre, University of Southampton & University Hospital Southampton NHS Foundation Trust

15Singapore Institute for Clinical Sciences (SICS), Agency for Science, Technology and Research (A\*STAR)

16Dr. D. Y. Patil Dental College & Hospital, Pune, Maharashtra, India

#corresponding author

Chin-Ying Stephen Hsu, DDS MS PhD

Faculty of Dentistry, National University of Singapore,

9 Lower Kent Ridge Rd, Singapore 119085

Email: denhsus@nus.edu.sg

**APPENDIX**

**Materials and Methods**

As expected, missing data was encountered because of the longitudinal study design with multiple time-points. To overcome the constraint of biased/overestimated results that may arise as a result of missing data, multiple imputation was used to impute the missing covariates, as a sensitivity test. Prior to performing imputation, assumption for missing data to be missing at random (MAR) was verified. Upon verification, Multiple Imputation by Chained Equation (MICE) was used (White et al. 2011) for imputation of missing covariates. MICE has been shown to be less biased than standard imputation approaches while properly accounting for the uncertainty arising from imputation (White et al. 2011). The outcome was used for imputation of missing predictors (Moons et al. 2006). Briefly, this method handles missing data in 3 steps: (1) imputes missing data *m* times to produce *m* complete datasets; (2) analyses each dataset using the selected statistical procedure; and (3) produces an overall estimate obtained from *m* results using Rubin’s rules (Rubin 1987). Predictive mean matching (PMM) was used for imputation of values for all missing covariates and 40 datasets were created. It has the advantage of maintaining the underlying distribution of data by imputing a random value from the set of observed values whose predicted values are closest to the predicted value from a specified regression model.

**Results**

Appendix Table 1. Potential determinants for moderate-extensive caries lesions

|  |  |
| --- | --- |
| **Risk Predictors** | **Moderate-extensive lesions [N(%)] (ICDAS codes 3-6)** |
| **Year 2** | **Year 3** |
| **Caries-affected****(dmfs>0)** | **Caries-free****(dmfs=0)** | **Caries-affected****(dmfs>0)** | **Caries-free****(dmfs=0)** |
| ***Ethnicity*** | *p=0.30* | *p=0.004* |
| Chinese | 14 (4.8%) | 279 (95.2%) | 85 (20.8%) | 323 (79.2%) |
| Malay | 6 (3.9%) | 148 (96.1%) | 51 (26.4%) | 142 (73.6%) |
| Indian | 1 (1.1%) | 87 (98.9%) | 13 (10.8%) | 107 (89.2%) |
| ***Mother’s education***  | *p=0.61* | *p=0.02* |
| Primary and below | 1 (4.2%) | 23 (95.8%) | 9 (32.1%) | 19 (67.0%) |
| Secondary | 10 (5.0%) | 189 (95.0%) | 63 (24.1%) | 198 (75.9%) |
| Polytechnic | 3 (2.1%) | 138 (97.9%) | 39 (21.8%) | 140 (78.2%) |
| Bachelor’s degree and above | 7 (4.2%) | 158 (95.8%) | 36 (14.6%) | 211 (85.4%) |
| ***Father’s education*** | *p=0.80* | *p=0.016* |
| Primary and below | 1 (4.5%) | 21 (95.5%) | 13 (43.3%) | 17 (56.7%) |
| Secondary | 2 (2.2%) | 87 (97.8%) | 22 (21.0%) | 83 (79.0%) |
| Polytechnic | 7 (4.4%) | 153 (95.6%) | 47 (24.1%) | 148 (75.9%) |
| Bachelor’s degree and above | 6 (3.8%) | 154 (96.2%) | 37 (18.05%) | 168 (81.95%) |
| ***Mother’s occupation*** | *p=0.70* | *p=0.19* |
| Unemployed | 5 (3.0%) | 161 (97.0%) | 36 (17.1%) | 174 (82.9%) |
| Non-professional | 8 (4.7%) | 162 (95.3%) | 54 (24.1%) | 170 (75.0%) |
| Professional | 7 (3.8%) | 176 (96.2%) | 55 (20.3%) | 216 (79.7%) |
| ***Monthly household income*** | *p=0.10* | *p=0.05* |
| 0-1999 SGD | 6 (7.5%) | 74 (92.5%) | 33 (30.6%) | 75 (69.4%) |
| 2000-3999 SGD | 6 (3.8%) | 153 (96.2%) | 42 (20.9%) | 159 (79.1%) |
| 4000-5999 SGD | 2 (1.5%) | 134 (98.5%) | 31 (18.3%) | 138 (81.7%) |
| >6000 SGD | 7 (5.8%) | 113 (94.2%) | 35 (18.0%)  | 159 (82.0%) |
| ***Prenatal smoke exposure*** | *p=0.05* | *p=0.18* |
| Yes | 12 (6.1%) | 184 (93.9%) | 61 (23.4%) | 200 (76.6%) |
| No | 8 (2.7%) | 293 (97.3%) | 81 (19.1%) | 342 (80.9%) |
| ***History of chronic maternal illness*** | *p=0.52* | *p=0.08* |
| Yes | 6 (3.0%) | 196 (97.0%) | 54 (17.5%) | 254 (82.5%) |
| No | 11 (4.1%) | 255 (95.9%) | 91 (22.9%) | 306 (77.1%) |
| ***In-between meal sweet snacks frequency*** | *p=0.43* | *p=0.01* |
| Twice or more per day | 6 (3.0%) | 192 (97.0%) | 61 (25.3%) | 180 (74.7%) |
| Once per day | 10 (5.0%) | 191 (95.0) | 54 (22.8%) | 183 (77.2%) |
| None | 2 (2.5%) | 78 (97.5%) | 12 (11.2%) | 95 (88.8%) |
| ***Parental reported cause for tooth decay*** | *p=0.23* | *p=0.20* |
| Sugar and bacteria | 9 (3.2%) | 276 (96.8%) | 68 (19.8%) | 276 (80.2%) |
| Sugar or bacteria | 0 (0.0%) | 15 (100.0%) | 4 (25.0%) | 12 (75.0%) |
| Ineffective toothbrushing | 1 (1.6%) | 61 (98.4%) | 16 (18.8%) | 69 (81.2%) |
| Tooth worms/heatiness | 8 (6.8%) | 109 (93.2%) | 39 (28.1%) | 100 (71.9%) |
| ***Night-time bottle feeding****¶* | *p=0.31* | *p=0.16* |
| Yes | 12 (3.4%) | 339 (96.6%) | 100 (22.9%) | 337 (77.1%) |
| No | 9 (5.1%) | 168 (94.9%) | 45 (18.4%) | 200 (81.6%) |
| ***Sharing feeding/drinking utensils*** | *p=0.34* | *p=0.13* |
| Frequently | 5 (4.8%) | 99 (95.2%) | 23 (19.2%) | 97 (80.8%) |
| Seldom | 5 (2.3%) | 209 (97.7%) | 65 (25.5%) | 190 (74.5%) |
| Never | 8 (4.9%) | 154 (95.1%) | 39 (18.4%) | 173 (81.6%) |
| ***History of allergies before 12 months (runny nose: 9-12 months; itchy rashes: 6-12 months)*** | *p<0.001* | *p=0.15* |
| Yes  | 6 (14.3%) | 36 (85.7%) | 17 (27.9%) | 44 (72.1%) |
| No | 12 (3.0%) | 390 (97.0%) | 120 (20.1%) | 478 (29.9%) |
| ***History of allergies before first tooth eruption (runny nose, itchy rashes)*** | *p=0.07* | *p=0.018* |
| Yes  | 10 (5.8%) | 161 (94.2%) | 63 (25.6%) | 183 (74.4%) |
| No | 8 (2.6%) | 301 (97.4%) | 86 (18.1%) | 389 (81.9%) |
| ***Intake of medications (oral medicines: 6-12 months; nebulizers: 9-12 months; antibiotics:9-12 months; medications for runny nose: 0-12 months)*** | *p=0.45* | *p=0.52* |
| Yes  | 6 (4.9%) | 116 (95.1%) | 41 (23.0%) | 137 (77.0%) |
| No | 10 (3.2%) | 305 (96.8%) | 100 (2.6%) | 385 (79.4%) |
| ***History of infection (ear infection/vomiting/diarrhoea) before first tooth eruption*** | *p=0.24* | *p=0.74* |
| Yes  | 7 (5.4%) | 122 (94.6%) | 39 (19.9%) | 157 (80.1%) |
| No | 11 (3.1%) | 340 (96.9%) | 110 (21.0%) | 414 (79.0%) |
| ***Mother’s brushing frequency*** | *p=0.94* | *p=0.012* |
| ≥3 times/day | 2 (3.4%) | 56 (96.6%) | 22 (28.9%) | 54 (71.0%) |
| 2 times/day | 14 (3.9%) | 343 (96.1%) | 98 (22.4%) | 340 (77.6%) |
| ≤1 times/day | 2 (3.1%) | 63 (96.9%) | 7 (9.6%) | 66 (90.4%) |
| ***Child has pre-existing medical condition*** | *p=0.98* | *p=0.06* |
| Yes | 4 (4.3%) | 90 (95.7%) | 36 (26.7%) | 99 (73.3%) |
| No | 17 (4.2%) | 388 (95.8%) | 102 (19.4%) | 423 (80.6%) |
| ***Maternal regular dental check-up*** | *p=0.85* | *p=0.52* |
| Yes | 8 (3.9%) | 195 (96.1%) | 52 (20.4%) | 203 (79.6%) |
| No | 10 (3.6%) | 267 (96.4%) | 75 (22.6%) | 257 (77.4%) |
| ***Presence of dry mouth at 18 months*** | *p=0.54* | *p=0.71* |
| Yes | 1 (6.7%) | 14 (93.3%) | 4 (23.5%) | 13 (76.5%) |
| No | 16 (3.7%) | 422 (96.3%) | 128 (19.9%) | 516 (80.1%) |
| ***Previous dental visit*** | *p=0.003* | *p=0.18* |
| Yes | 2 (22.2%) | 7 (77.8%) | 6 (33.3%) | 12 (66.7%) |
| No | 16 (3.4%) | 455 (96.6%) | 142 (20.5%) | 551 (79.5%) |

*¶up to 12 months for ECC at 2 years outcome and up to 24 months for ECC at 3 years outcome*

Appendix Table 2. Medical CRA models predicting caries risk at 2 years –Risk determinants

|  |
| --- |
| **Model I-A: All lesions (ICDAS codes 2-6)** |
| **Significant risk predictors** | **N** | **Adjusted\* OR (95%CI)** | **p-value** |
| ***Monthly household income*** |  |  |  |
| 0-1999 SGD | 42 | 0.28 (0.07-1.06) | 0.06 |
| 2000-3999 SGD | 92 | 0.26 (0.08-0.79) | 0.017 |
| 4000-5999 SGD | 74 | 0.56 (0.22-1.38) | 0.208 |
| ≥6000 SGD | 68 | 1 |  |
| ***History of chronic maternal illness*** |  |  |  |
| Yes  | 115 | 0.16 (0.07-0.38) | <0.001 |
| No | 161 | 1 |  |
| ***History of allergies before first tooth eruption (runny nose, itchy rashes)*** |  |  |  |
| Yes  | 107 | 2.58 (1.44-6.14) | 0.014 |
| No | 169 | 1 |  |
| ***Intake of medications (oral medicines: 6-12 months; nebulizers: 9-12 months; antibiotics:9-12 months; medications for runny nose: 0-12 months)*** |  |  |  |
| Yes  | 81 | 2.98 (1.44-6.14) | 0.003 |
| No | 195 | 1 |  |
| ***History of infection (ear infection/vomiting/diarrhoea) before first tooth eruption*** |  |  |  |
| Yes  | 75 | 3.11 (1.44-6.74) | 0.004 |
| No | 201 | 1 |  |
|  |
| **Model I-B: Moderate-extensive lesions (ICDAS codes 3-6)** |
| **Significant risk predictors** | **N** | **Adjusted# OR (95%CI)** | **p-value** |
| ***Maternal childbearing age (in years)*** | 299 | 1.26 (1.00-1.48) | 0.04 |
| ***Prenatal smoke exposure*** |  |  |  |
| Yes  | 118 | 6.79 (1.40-32.90) | 0.017 |
| No | 181 | 1 |  |
| ***History of allergies before 12 months (runny nose: 9-12 months; itchy rashes: 6-12 months)*** |  |  |  |
| Yes  | 24 | 9.18 (1.34-62.73) | 0.024 |
| No | 275 | 1 |  |

\*Risk predictors (17) in Model I-A: ***Prenatal factors:*** Ethnicity, maternal childbearing age, maternal occupation, parity status, history of chronic maternal illness, monthly household income, prenatal smoke exposure; ***Peri-natal factor***: mode of delivery; ***Post-natal factors:*** History of allergies before 12 months (runny nose: 9-12 months; itchy rashes: 6-12 months), History of allergies before first tooth eruption (runny nose, itchy rashes), Intake of medications (oral medicines: 6-12 months; nebulizers: 9-12 months; antibiotics: 9-12 months; medications for runny nose: 0-12 months), BMI (12M), History of infection (ear infection/vomiting/diarrhoea) before first tooth eruption, number of teeth present (12M), duration of breastfeeding (in months), child’s age for first tooth eruption, age of introduction of solid food (in months)

**#**Model I-B: Having the same predictors as those of Model I-A, but without monthly household income and mode of delivery (owing to high standard error)

 *§Accuracy calculated as [(sensitivity x prevalence of outcome) + (specificity x non-prevalence of outcome)/100].*

Appendix Table 3. Medical CRA models predicting caries risk at 3 years –Risk determinants

|  |
| --- |
| **Model II-A: All lesions (ICDAS codes 2-6)** |
| **Significant risk predictors** | **N** | **Adjusted OR (95%CI)** | **p-value** |
| ***Ethnicity*** |  |  |  |
| Chinese | 173 | 2.67 (1.00-7.10) | 0.048 |
| Malay | 99 | 2.67 (0.95-7.56) | 0.06 |
| Indian | 40 | 1 |  |
| ***Mother’s education***  |  |  |  |
| Primary or below | 8 | 0.50 (0.07-3.59) | 0.49 |
| Secondary | 115 | 0.33 (0.12-0.91) | 0.032 |
| Pre-university | 93 | 0.52 (0.23-1.18) | 0.120 |
| University and above | 96 | 1 |  |
| ***Mother’s occupation*** |  |  |  |
| Unemployed/retired/homemaker/student | 88 | 0.67 (0.34-1.32) | 0.252 |
| Non-professional | 93 | 2.50 (1.30-4.83) | 0.006 |
| Professional | 131 | 1 |  |
| ***History of chronic maternal illness*** |  |  |  |
| Yes  | 134 | 0.34 (0.19-0.62) | <0.001 |
| No | 178 | 1 |  |
| ***Prenatal smoke exposure*** |  |  |  |
| Yes  | 116 | 2.57 (1.33-4.99) | 0.005 |
| No | 196 | 1 |  |
| ***Parental reported cause for tooth decay*** |  |  |  |
| Sugar and bacteria | 187 | 0.50 (0.25-0.99) | 0.049 |
| Sugar or bacteria | 5 | 7.40 (0.57-94.77) | 0.124 |
| Ineffective toothbrushing | 51 | 0.49 (0.20-1.19) | 0.117 |
| Tooth worms/heatiness | 69 | 1 |  |
| ***Mother’s brushing frequency*** |  |  |  |
| ≥3 times/day | 38 | 3.58 (1.12-11.42) | 0.031 |
| 2 times/day | 230 | 2.79 (1.16-6.70) | 0.021 |
| ≤1 times/day | 44 | 1 |  |
| ***In-between meal sweet snacks frequency*** |  |  |  |
| ≥2 times/day | 123 | 3.21 (1.50-6.95) | 0.003 |
| 1 times/day | 124 | 0.93 (0.44-1.97) | 0.855 |
| None | 65 | 1 |  |
| ***Use of pacifier (in months)*** | 312 | 0.96 (0.93-0.99) | 0.027 |
| ***Number of teeth erupted at 12 months*** | 312 | 1.16 (1.00-1.34) | 0.047 |
| ***Child’s age for first tooth eruption (in months)*** | 312 | 1.19 (1.02-1.40) | 0.03 |
|  |
| **Model II-B: Moderate-extensive lesions (ICDAS codes 3-6)** |
| **Significant risk predictors** | **N** | **Adjusted OR (95%CI)** | **p-value** |
| ***Ethnicity*** |  |  |  |
| Chinese | 173 | 5.09 (1.38-18.69) | 0.014 |
| Malay | 99 | 6.16 (1.60-23.76) | 0.008 |
| Indian | 40 | 1 |  |
| ***History of chronic maternal illness*** |  |  |  |
| Yes  | 134 | 0.48 (0.24-0.94) | 0.033 |
| No | 178 | 1 |  |
| ***Mother’s brushing frequency*** |  |  |  |
| ≥3 times/day | 38 | 4.03 (0.91-17.85) | 0.066 |
| 2 times/day | 230 | 3.74 (1.09-12.79) | 0.035 |
| ≤1 times/day | 44 | 1 |  |
| ***In-between meal sweet snacks frequency*** |  |  |  |
| ≥2 times/day | 123 | 3.60 (1.42-9.14) | 0.007 |
| 1 times/day | 124 | 1.19 (0.47-3.05) | 0.707 |
| None | 65 | 1 |  |
| ***Child has pre-existing medical condition*** |  |  |  |
| Yes | 68 | 2.24 (1.08-4.67) | 0.031 |
| No | 244 | 1 |  |
| ***Use of pacifier (in months)*** | 312 | 0.94 (0.90-0.98) | 0.003 |

\*Risk predictors (22) in Model II-A/B: ***Prenatal factors:*** Ethnicity, Household income, Maternal educational levels, Maternal occupation, Father educational levels, Parity status, History of chronic maternal illness, Prenatal smoke exposure; ***Post-natal factors:*** Parental belief for cause of tooth decay, Mother’s brushing frequency, Child’s pre-existing medical condition, Frequency of in between meal sweet snacks, duration of pacifier use (in months), duration of breastfeeding (in months), maternal regular dental check-up, child’s brushing frequency, number of teeth (M12), sharing feeding/drinking utensils, night-time bottle feeding, child’s age for first tooth eruption (in months), dry mouth (M18), previous dental visit

*§Accuracy calculated as [(sensitivity x prevalence of outcome) + (specificity x non-prevalence of outcome)/100].*

Appendix Table 4. Medical CRA model for 1-year increase in caries (∆M36-24 dmfs>0)

|  |
| --- |
| **Model II-A (for outcome ∆M36-24 dmfs>0)** |
| **Adjusted\* N** | **AUC (95%CI)** | **Cut-off** | **Sensitivity** | **Specificity** | **Accuracy** | **Pseudo R2** |
| 251 | 0.78 (0.73-0.84) | 0.44 | 74 | 72 | 73 | 0.31 |
| **Model II-B (for outcome ∆M36-24 dmfs>0)** |
| **Adjusted\* N** | **AUC (95%CI)** | **Cut-off** | **Sensitivity** | **Specificity** | **Accuracy** | **Pseudo R2** |
| 251 | 0.80 (0.74-0.86) | 0.260 | 73 | 70 | 71 | 0.33 |

\*Risk predictors (22): ***Prenatal factors:*** Ethnicity, household income, maternal educational levels, maternal occupation, father educational levels, parity status, history of chronic maternal illness, prenatal smoke exposure; ***Post-natal factors:*** Parental belief for cause of tooth decay (24 months), mother’s brushing frequency (24 months), child’s pre-existing medical condition (6 months), frequency of in between meal sweet snacks (24 months), duration of pacifier use (in months), duration of breastfeeding (in months), maternal regular dental check-up (24 months), child’s brushing frequency (24 months), number of teeth (12 months), sharing feeding/drinking utensils, night-time bottle feeding (24 months), child’s age for first tooth eruption (in months), dry mouth (18 months), previous dental visit (24 months).

Appendix Table 5. Internal validation of medical CRA models

|  |  |
| --- | --- |
| **Medical CRA model for Year 2** | **Medical CRA model for Year 3** |
| **Training set (70%)****(N = 368)** | **Validation set (30%)****(N = 167)** | **Training set (70%)****(N = 503)** | **Validation set (30%)** **(N = 218)** |
| **Model I-A** | **Model II-A** |
| **AUC (95%CI)** | **AUC (95%CI)** | **AUC (95%CI)** | **AUC (95%CI)** |
| 0.76 (0.69-0.82) | 0.68 (0.58-0.79) | 0.72 (0.67-0.77) | 0.64 (0.56-0.72) |
| **Model I-B** | **Model II-B** |
| **AUC (95%CI)** | **AUC (95%CI)** | **AUC (95%CI)** | **AUC (95%CI)** |
| 0.83 (0.72-0.94) | 0.66 (0.45-0.87) | 0.76 (0.70-0.82) | 0.67 (0.58-0.77) |

**REFERENCES**

Moons KG, Donders RA, Stijnen T, Harrell FE, Jr. 2006. Using the outcome for imputation of missing predictor values was preferred. J Clin Epidemiol. 59(10):1092-1101.

Rubin. 1987. Multiple imputation for nonresponse in surveys. New York: Wiley.

White IR, Royston P, Wood AM. 2011. Multiple imputation using chained equations: Issues and guidance for practice. Stat Med. 30(4):377-399.