Validity of Caries Risk Assessment Programs in Preschool Children

Xiao-Li Gao[1], Di Wu[2], Edward Chin Man Lo[1], Chun Hung Chu[3], Stephen Chin-Ying Hsu[4], May Chun Mei Wong[1]

[1] Dental Public Health, Faculty of Dentistry, The University of Hong Kong, 34 Hospital Road, Hong Kong

[2]

[3] Oral Diagnosis and Polyclinics, Faculty of Dentistry, The University of Hong Kong, 34 Hospital Road, Hong Kong

[4] Faculty of Dentistry, National University of Singapore, 11 Lower Kent Ridge Road, Singapore 119083

Keywords: *early childhood caries, multifactorial modelling, caries prediction, sensitivity, specificity, prospective study*

Correspondence to : Dr. Gao Xiaoli
Research Assistant Professor
3/F, Periodontology and Dental Public Health,
Prince Philip Dental Hospital,
34 Hospital Road,
Sai Ying Pun, Hong Kong

Tel: (852) 2859 0481
Fax: (852) 2858 7874
Email: gaoxl@hkucc.hku.hk

Number of words in the Abstract: 200
Total words (from Abstract to Acknowledgements): 2698
Number of tables and figures: 4
Number of cited references: 25

Validity of Caries Risk Assessment Programs in Preschool Children

ABSTRACT

Identifying caries-prone children through caries risk assessment (CRA) contributes to targeted prevention and evidence-based treatment planning. This study aimed to compare the validity of various CRA programs (CAT, CAMBRA, Cariogram, and NUS-CRA), without biological tests (screening mode) and with biological tests (comprehensive mode), in predicting early childhood caries. A total of 544 children aged 3 years underwent oral examination and biological tests (saliva flow, buffering capacity and abundance of cariogenic bacteria). Their parents completed a questionnaire. Children's caries risk was predicted using the study programs. After 12 months, caries increment in 485 (89%) children was compared with the baseline risk predictions. Reasoning-based programs (CAT and CAMBRA screening) had high sensitivity (≥93.8%) but low specificity (<=43.6%) in predicting caries. CAMBRA comprehensive assessment reached a better balance (sensitivity/specificity of 83.7%/62.9%). Algorism-based programs (Cariogram and NUS-CRA) generated better predictions. The sensitivity/specificity of NUS-CRA screening and comprehensive models were 73.6%/84.7% and 86.5%/82.4%, respectively, higher than those of the Cariogram screening (68.0%/75.9%) and comprehensive assessment (79.2%/65.5%). NUS-CRA comprehensive model met the criteria for a useful CRA tool (sensitivity+specificity ⩾160%), while its screening model approached that target. Findings of this study supported algorism-based approach and the usefulness of NUS-CRA in identifying children susceptible to caries.

INTRODUCTION

As a common childhood disease, dental caries imposes significant threat on the wellbeing of preschool children and constitutes a significant financial burden on families and society (US/DHHS, 2007). Recent epidemiological evidence revealed a "silent epidemic" of early childhood caries (ECC) and its polarized distribution - the main caries burden is confined to a minority of high-risk children (Seppa, 2001; US/DHHS, 2007). Such a disease pattern has steered tremendous efforts into developing caries risk assessment (CRA) programs to identify susceptible children for early prevention and intervention so that targeted caries control and optimized treatment planning can be achieved (NIH consensus panel, 2001).

Several multifactorial CRA programs have been developed and made applicable to preschool children, including the Caries-risk Assessment Tool (CAT) (AAPD, 2006), the Caries Management by Risk Assessment (CAMBRA) program (Ramos-Gomez et al., 2007), the Cariogram (Bratthall & Petersson, 2005), and the NUS-CRA biopsychosocial models (Gao et al., 2010). Under each program, caries risk can be estimated based on a questionnaire and clinical observations (screening assessment) and further refined with the aid of salivary and/or microbiological tests (comprehensive assessment). Despite the availability of these CRA programs, their validity in predicting dental caries remains largely unknown let alone their clinical potentials (Tellez et al., 2012).

The validity of a CRA program is often measured by its sensitivity (Se) (proportion of high-risk people who are classified as such) and specificity (Sp) (proportion of low-risk people who are classified correctly). For a CRA tool to be practically useful, a common consensus is that it should reach a sum of Se and Sp (Se+Sp) of at least 160%, and, ideally, a sound balance between these two parameters (Stamm, 1988, Zero et al., 2001). Current reports on the Se/Sp

of above-mentioned CRA programs were scarce, especially those from prospective studies. Although risk-based intervention guided by CAMBRA appeared to be effective in controlling caries in a group of adults (Featherstone et al., 2005), no attempt was made to evaluate its accuracy in assessing caries risk hence how well the intervention was "risk-based". Validity of CAT was investigated in a cross-sectional study among toddlers, where high Se (100%) and low Sp (3%) were found (Yoon et al., 2012); when socioeconomic factors were excluded, a better balance was observed (Se/Sp=86%/69%). The reported Se/Sp of Cariogram varied from 66%/60% (Utreja et al., 2010), 73%/60% (Petersson et al., 2010), to 83%/85% (Campus et al., 2012) among schoolchildren and from 46%/88% (Holgerson et al., 2009) to 71%/66% (Gao et al., 2010) in preschoolers. The NUS-CRA program has been tested among Singaporean preschoolers. The Se/Sp of its screening and comprehensive models were 82%/73% (Gao et al., 2010) and 81%/85% (Gao et al., under review), respectively. Nevertheless, it is yet to be tested in other populations.

Since the current evidence on the validity of these CRA programs (CAT, CAMBRA, Cariogram, and NUS-CRA) is far from being complete, this study aimed to simultaneously compare their validity in a preschool child population.

MATERIALS & METHODS

Participants

For evaluating the validity of the CRA programs, with an expected Se/Sp of 80%/80%, an estimated "% with new caries (Δdmft>0) in 12 months" of 35% (based on our previous data), a targeted precision level of 6%, and a confidence level of 95%, 489 subjects were needed. Allowing an attrition rate of 15%, a sample size of 575 was considered as appropriate.

Ethical approval was obtained from the Institutional Review Board of the University of Hong Kong (#08-400). With parental written consents, participants were recruited from four kindergartens located in different districts of Hong Kong. All kindergarten grade-1 children (3 years of age) were eligible to participate, except for those who were unable to cooperate in the related procedures or who had severe medical conditions.

Baseline Data Collection

1. Questionnaire

A pre-tested questionnaire was completed by parents to gather information on their socio-economic status, child's demographic background, parental knowledge on and attitude towards oral health, child's oral health habits (infant feeding history, diet, oral hygiene, use of topical fluorides, and dental attendance) and child's systemic health and medication.

2. Dental examination

All children were examined by a dentist who was trained and calibrated against an experienced oral epidemiologist. A satisfactory reliability (>90%) was achieved before the completion of the training session. Duplicate examinations were carried out on 10% of participants for assessing the intra-examiner reliability.

Children were examined in a supine position, with an illuminated mouth mirror and CPI probe. The tooth status was assessed by visual inspection, aided by tactile inspection if necessary. No radiographs were taken. Dental caries was registered at the cavitation level according to the World Health Organization criteria (WHO, 1997). White-spot lesion, which is a risk indicator in CAT and CAMBRA program, was also recorded, together with any developmental defect (e.g. hypoplasia) or dental appliance. The oral hygiene status was

evaluated using Silness-Löe Plaque Index (PI) (Silness & Löe, 1964).

3. Salivary and microbiological tests

Stimulated saliva flow rate, saliva buffering capacity, and levels of *mutan* Streptococci and *Lactobacilli* were evaluated, using Dentobuff®, Dentocult® MS Strip *mutans*, and Dentocult® LB kits (Orion Diagnostica, Finland). Standard laboratory procedures were followed for the incubation of bacteria, acquirement of readings, and disposal of biological wastes.

Caries Risk Assessment

Children's caries risk was assessed by using the study programs (Table 1). The algorithms of NUS-CRA models were built through epidemiological data collected from preschoolers in Singapore. Other programs were developed based on known literature. With NUS-CRA and Cariogram, children's caries risk was calculated through algorithms and expressed as "% chance of caries", which was further classified into 5 risk groups. Under CAT and CAMBRA, children were classified into 3 risk groups through manual charting using a checklist of important factors/indicators.

Rating criteria stipulated in the user instructions of each program were followed. In view of the recent findings on CAT (Yoon et al., 2012), children's caries risk was also assessed after excluding socioeconomic factors from CAT. Due to the constraints of the fieldwork in kindergartens, data on caries experience and bacterial levels were only collected in children, although CAT and CAMBRA suggest such data to be collected also from mothers/caregivers and/or siblings.

Follow-up of Disease Outcome

After 12 months, children's tooth status was re-evaluated by the same examiner following the above-mentioned method. The examiner was masked to the risk predictions generated from the study programs.

Data Analysis

The data were analyzed with the Statistical Package for Social Sciences (version 20). Caries increment among children who were categorized into different risk groups under various programs was compared using statistical methods as appropriate (denoted in table footnote). To evaluate the validity of each CRA program in predicting "new caries" (Δdmft>0)", the calculation of Se/Sp requires the determination of a cut-off point to categorize children dichotomously into "susceptible" and "non-susceptible". With CAT and CAMBRA, there were two possible cut-off points ("≥moderate risk" and "≥high risk"). For Cariogram and NUS-CRA whose risk prediction is in a continuous scale (% chance), Receiver Operation Characteristics (ROC) curve was plotted to identify the optimal cut-off point that generated the highest Se+Sp. The Area Under Curve (AUC) value was also calculated to reflect the overall performance across all possible cut-off points. The significant level was set at 0.05.

RESULTS

Among 585 children in the 4 kindergartens, 544 were examined at baseline. After 12 months, 485 (89%) children (261 boys and 224 girls) were followed up. The intra-examiner reliability was high (Kappa=0.964). No significant difference was found between children who completed the study and those lost to follow-up in their socio-demographic background and baseline caries experience, except that more girls than boys did not complete the study (14.5% vs. 7.4%; $p<0.05$).

Within 12 months, 178 (36.7%) children developed new caries (Δdmft>0) (Table 2). The

mean (SD) increment in dmft was 0.78 (1.36). With CAT and CAMBRA, the majority of children were considered as “high risk”; only a small proportion was defined as “moderate risk”. Under CAT, no participant was rated as “low risk”. In contrast, under Cariogram and NUS-CRA, the majority of children were defined as “very low” or “low” risk. Overall, children’s new caries increased from lower to higher risk groups under all programs, as observed from the “mean caries increment” (Δdmft) and/or “% with new caries” (Δdmft>0). Nevertheless, no statistically significant difference in caries increment was observed between some of the risk groups.

To assess the Se/Sp of the programs in predicting caries, both possible cut-off points (“≥moderate risk” and “≥high risk”) were explored for CAMBRA (Table 3). With CAT, since no child was considered as “low risk”, only one cut-off point (“≥high risk”) could be used. For Cariogram and NUS-CRA, the best cut-off points identified in the ROC analysis were selected. Across all programs, “susceptible” children (predicted risk above the cut-off point) consistently had higher “mean caries increment” and “% with new caries” than “non-susceptible” children (all $p<0.05$). For CAMBRA, compared with “≥moderate risk”, “≥high risk” appeared to be a better cut-off point generating higher Se+Sp. CAT had extremely high Se (>98%) but low Sp (<6%), under both screen and comprehensive assessments including or excluding socioeconomic risk factors. CAMBRA screening assessment also had a high Se (93.8%) and low Sp (43.6%); its comprehensive assessment reached a relatively balanced Se (83.7%) and Sp (62.9%). Cariogram screening assessment was less sensitive (68.0% vs. 79.2%) but more specific (75.9% vs. 65.5%) than its comprehensive version, with a similar Se+Sp (144% vs. 145%). Compared with its screening version, NUS-CRA comprehensive assessment had a higher sensitivity (86.5% vs. 73.6%), similar specificity (82.4% vs. 84.7%), and higher Se+Sp (169% vs. 158%). Among all models, only NUS-CRA comprehensive assessment reached a

Se+Sp above 160%, even after excluding "past caries", which is regarded as the strongest caries indicator (Zero *et al*., 2001). NUS-CRA screening model, with a Se+Sp of 158%, approached that target.

The ROC analysis showed that both the screening and comprehensive versions of NUS-CRA generated better prediction (higher AUC) than their Cariogram counterparts (Figure 1). Nevertheless, for both program, no significant difference was found in the AUC of their screening and comprehensive assessments.

DISCUSSION

The concept of risk-based dental management is not new to professionals. Its clinical application however remains limited. Feedback from clinicians shows both promises and challenges (Doméjean-Orliaguet et al., 2006; Nainar & Straffon, 2006; Gonzalez & Okunseri, 2010). In this connection, collecting scientific evidence on the validity of existing CRA programs may shed light on their potentials in dental practice and inform possible improvements.

The programs tested in this study represent two types of risk assessment tools, namely reasoning-based (CAT and CAMBRA) and algorism-driven programs (Cariogram and NUS-CRA). In the former, important risk factors and indicators were synthesized into a checklist or guideline and one's risk to disease is qualitatively estimated, whereas in the latter, one's risk was quantitatively calculated through algorithms. Our findings support a superior validity of algorism-driven programs, as evidenced by the higher Se+Sp of Cariogram and NUS-CRA, compared with CAT and CAMBRA. Algorism-driven approach was also successfully used in

the medical field for predicting chronic diseases, such as cardiovascular disease (e.g. Framingham risk equation, QRISK, and ASSIGN) (Dent, 2010).

Undoubtedly, reasoning-based programs such as CAT and CAMBRA are useful pedagogical tools for explaining the caries etiology and dynamics (Featherstone, 2004) to dental students and novice dentists. However, since experienced dentists often can estimate patients' caries risk to certain accuracy (60-70%) (Saemundsson, 1996), a CRA program is useful only if it significantly improves dentists' clinical judgment. Our results reveal a high Se but low Sp (i.e. a high false positive rate) of CAT and CAMBRA. This echoes the finding of a recent study on CAT (Yoon et al., 2012). Such overestimation may have stemmed from the classification criteria, under which, some single factors/indicators alone is sufficient to justify a "high risk" diagnosis. One may question that, without collecting data from mothers/caregivers and siblings, we might have underestimated the validity of CAT and CAMBRA. A careful scrutiny into the rating criteria would however suggest that adding these factors into the assessment may further exacerbate the overestimation of children's caries risk. In addition, collecting data on mothers/caregivers and siblings' caries experience and bacterial count may often be impractical to clinicians and public health workers.

With high sensitivity, CAT and CAMBRA may be useful under some clinical scenarios where the failure to identify and treat any high-risk child is absolutely unbearable and is the only concern. Nevertheless, the low Sp thus overestimation of risk leads to overtreatment and waste of resources. Fine-tuning some of the rating criteria may be necessary for improving the performance of these programs. In a group of children from low-income families, excluding socioeconomic risk factors from the CAT assessment alleviated the overestimation on caries risk (Yoon et al., 2012). Nevertheless, such improvement was minimal in our study among

children from diverse socioeconomic backgrounds, where misclassified cases due to low socioeconomic status alone were not many.

Screening models of Cariogram and NUS-CRA reached reasonably accuracy, supporting the possibility of simple and inexpensive CRA. Although saliva flow rate and buffering capacity were included in Cariogram assessment, NUS-CRA models excluded them, since they failed to be useful caries predictors due to inaccurate measurement of saliva flow and rare salivary abnormalities in young children (Gao et al., 2010). The inclusion of biological tests in Cariogram and NUS-CRA further improved their sensitivity, while for CAMBRA, the improvement was in its specificity (i.e. low false positive rate). This is expected since several risk factors are required to justify a higher risk prediction when low bacterial levels are detected, as stipulated in CAMBRA (Ramos-Gomez et al., 2007).

Unlike other programs, which either target a specific age group or provide different versions for various age groups, Cariogram is intended to provide CRA to all ages (Bratthall & Petersson, 2005). However, the current evidence from this and other studies (Holgerson et al., 2009; Gao et al., 2010) unanimously pointed to an unsatisfactory performance of Cariogram in preschool children. Since unique risk factors may be involved in the occurrence of ECC, it may be reasonable to incorporate some age-specific factors into Cariogram and recalibrate the built-in algorisms for its better applicability to young children.

While the field of CRA development continues to evolve, our findings supported the algorism-based modelling and the contribution of some biological tests for ECC prediction. A stable Se/Sp of NUS-CRA was observed in this sample and Singaporean children (Gao et al., 2010; Gao et al., under review), suggesting the usefulness of this program in oriental

populations with fluoridated water supply. Studies in non-fluoridated communities may further shed light on the robustness of this program across different populations. Echoing our previous findings (Gao et al., under review), this study showed that NUS-CRA comprehensive version is not compromised even after excluding "past caries" from the model thus enables early prediction before caries occurs on any tooth of the child's mouth.

Acknowledgement

We appreciate the support of the participating kindergartens. This study was financially supported by the Health and Health Service Research Fund of the Hong Kong SAR government (project #07080741). The authors declare no conflict of interest.

References

American Academy of Pediatric Dentistry (AAPD) (2006). Policy on use of a caries-risk assessment tool (CAT) for infants, children, and adolescents. Reference Manual V 30/No7 08 / 09.

Bratthall D, Hansel Petersson G (2005). Cariogram - a multifactorial risk assessment model for a multifactorial disease. Community Dent Oral Epidemiol 33:256-64.

Campus G, Cagetti MG, Sale S, Carta G, Lingström P. Cariogram validity in schoolchildren: a two-year follow-up study. Caries Res. 2012;46(1):16-22.

Dent TH (2010). Predicting the risk of coronary heart disease I. The use of conventional risk markers. Atherosclerosis 213(2):345-51.

Doméjean-Orliaguet S, Gansky SA, Featherstone JD (2006). Caries risk assessment in an educational environment. J Dent Educ 70(12):1346-54.

Featherstone JDB, Gansky SA, Hoover CI, Rapozo-Hilo M, Weintraub JA, Wilson RS, et al (2005). A randomized clinical trial of caries management by risk assessment. Caries Res 39:295(Abstr #25).

Featherstone JD (2004). The caries balance: the basis for caries management by risk assessment. Oral Health Prev Dent 2 Suppl 1:259-64.

Gao XL, Hsu CY, Xu Y, Hwarng HB, Loh T, Koh D (2010). Building caries risk assessment models for children. J Dent Res 89(6):637-43.

Gao XL, Hsu CYS, Xu Y, Loh T, Koh D, Hwarng HB (under review). Role of Microbiological Factors in Predicting Early Childhood Caries. J Dent Res.

Gonzalez CD, Okunseri C (2010). Senior dental students' experience with Cariogram in a pediatric dentistry clinic. J Dent Educ 74(2):123-9.

Holgerson PL, Twetman S, Stecksen-Blicks C (2009). Validation of an agemodified caries risk assessment program (Cariogram) in preschool children. Acta Odontol Scand 67:106-112.

Nainar SM, Straffon LH (2006). Predoctoral dental student evaluation of American Academy of Pediatric Dentistry's Caries-Risk Assessment Tool. J Dent Educ 70(3):292-5.

National Institute of Health (NIH) consensus panel (2001). National Institute of Health consensus development conference statement. Presented at the Consensus Development Conference on Diagnosis and Management of Dental Caries Throughout Life, March 26-28, 2001. Natcher Conference Center, National Institutes of Health, Bethesda, MD. USA.

Petersson GH, Isberg PE, Twetman S (2010). Caries risk assessment in school children using a reduced Cariogram model without saliva tests. BMC Oral Health 19;10:5.

Ramos-Gomez FJ, Crall J, Gansky SA, Slayton RL, Featherstone JD (2007). Caries risk assessment appropriate for the age 1 visit (infants and toddlers). J Calif Dent Assoc 35(10):687-702.

Saemundsson S (1996). Dental caries prediction by clinicians and neural networks. Imprint: Ann Arbor, Mich. University Microfilms International. Thesis (Ph.D.) -University of North Carolina at Chapel Hill.

Seppa L (2001). The future of preventive programs in countries with different systems for dental care. Caries Res 35 (Suppl 1):26-9.

Silness J, Löe H (1964). Periodontal disease in pregnancy. II. Correlation between oral hygiene and periodontal condition. Acta Odontol Scand 22:121-35.

Stamm J, Disney J, Graves R, Bohannan H, Abernathy J (1988). The University of North Carolina Caries Risk Assessment Study. I: Rationale and content. J Public Health Dent 48:225-232.

Tellez M, Gomez J, Pretty I, Ellwood R, Ismail A (2012). Evidence on existing caries risk assessment systems: are they predictive of future caries? Community Dent Oral Epidemiol. doi: 10.1111/cdoe.12003.

U.S. Department of Health and Human Services (DHHS) (2007). Trends in oral health status: United States, 1988-1994 and 1999-2004. Vital Health Stat 11 (248):1-92.

Utreja D, Simratvir M, Kaur A, Kwatra KS, Singh P, Dua V. An evaluation of the Cariogram as a predictor model. Int Dent J. 2010 Aug;60(4):282-4.

World Health Organization (WHO) (1997). Oral health surveys: Basic methods. 4th edition. WHO Geneva.

Yoon RK, Smaldone AM, Edelstein BL (2012). Early childhood caries screening tools: a comparison of four approaches. J Am Dent Assoc 143(7):756-63.

Zero D, Fontana M, Lennon AM (2001). Clinical applications and outcomes of using indicators of risk in caries management. J Dent Educ 65:1126-1132.