Diagnostic test accuracy using digital retinal imaging in the detection of any diabetic retinopathy by graders in Vietnam, against a reference standard from the UK

Curran K1, Congdon N2,3, Hoang TT4,5, Lohfeld L1,6, Nguyen H7, Nguyen Q8, Nguyen V9, Dardis C9, Tran H7, Huong TH10, Vu AT10, Tung MQ6, Peto T1,9

Affiliations: 1Centre for Public Health, Queen's University Belfast 2Orbis International, NY, 3Zhongshan Ophthalmic Centre, Guangzhou, China 4Ophthalmology Department, Hanoi Medical University, Vietnam, 5Save Sight Institute, The University of Sydney School of Medicine, Australia 6Eye Hospital of Wenzhou Medical University, Wenzhou, Zhejiang Province, China 7Orbis International, Vietnam 8Vitreo-Retina Department, Ho Chi Minh Eye Hospital, Vietnam 9Ophthalmology Department, Belfast Health and Social Care Trust 10Community Department, Vietnam National Institute of Ophthalmology, Hanoi, Vietnam

PURPOSE: To compare the diagnostic test accuracy (DTA) of trained diabetic retinopathy (DR) graders in Vietnam against a reference standard from the UK, and assess the impact of supplementary grader training.

METHODS

DR grader training (Mar 2018): Level 1 DR graders (nurses, endocrinologists, and general practitioners), level 2 graders (mid-level ophthalmologists) and level 3 graders (senior ophthalmologists) in Vietnam were trained to grade DR severity by UK graders.

Phase 1 (Sept-Nov 2018): Fundus images were graded for DR by level 1, 2 and 3 graders in Vietnam. A trained grader from the UK graded all images retrospectively in masked fashion and served as the reference standard.

Analysis: DTA was calculated using sensitivity, specificity, and positive and negative predictive values (with 95% CIs).

Phase 2 (Mar-Apr 2021): An additional subset of images were graded by level 1, 2 and 3 graders in Vietnam. A trained grader from the UK graded all images retrospectively in masked fashion and served as the reference standard.

RESULTS

• Phase 1: The sensitivity for detecting ‘any DR’ was low among all graders in Vietnam (level 1 graders; 41.8 (34.0, 50.0), level 2 graders; (42.5 (34.6, 50.7), level 3 graders; 42.2 (33.2, 51.8).

• Phase 2: The sensitivity for detecting ‘any DR’ improved drastically for all graders in Vietnam after remedial test and training was provided (level 1 graders; 95.8 (97.3, 98.9), level 2 graders; (93.0 (83.7, 97.4), level 3 graders; 85.4 (70.1, 93.9).

• Any DR was defined as grades (R1-R3a) based on the UK DR Classification System.

CONCLUSION

• Grading accuracy was low in Vietnam in the first six months of implementing a training programme but after further training was delivered, DTA among all graders increased.

• Findings from this study are particularly important for DR programme planners in low- and middle-income countries.