**Supplementary materials**

**Near-infrared spectroscopy as an alternative method** **to assess** **typhaneoside and** **isorhamnetin-3-O-glucoside in different processed products of Pollen Typhae**

Running head: rapid quantitative analysis of typhaneoside and isorhamnetin-3-O-glucoside

Fei Suna,b,c,#, Weijun Zhaoa,#, Kaiyang Wanga, Shumei Wanga,b,c , Shengwang Lianga,b,c,\*

a Department of Traditional Chinese Medicine Analysis, School of Traditional Chinese Medicine, Guangdong Pharmaceutical University, Guangzhou 510006, PR China

b Key Laboratory of Digital Quality Evaluation of Chinese Materia Medica of State Administration of Traditional Chinese Medicine, Guangzhou 510006, PR China

c Guangdong Academies Traditional Chinese Medicine Quality Engineering Technology Research Center, Guangzhou 510006, PR China

# These authors contributed equally to this study and share first authorship.

\* Correspondences:

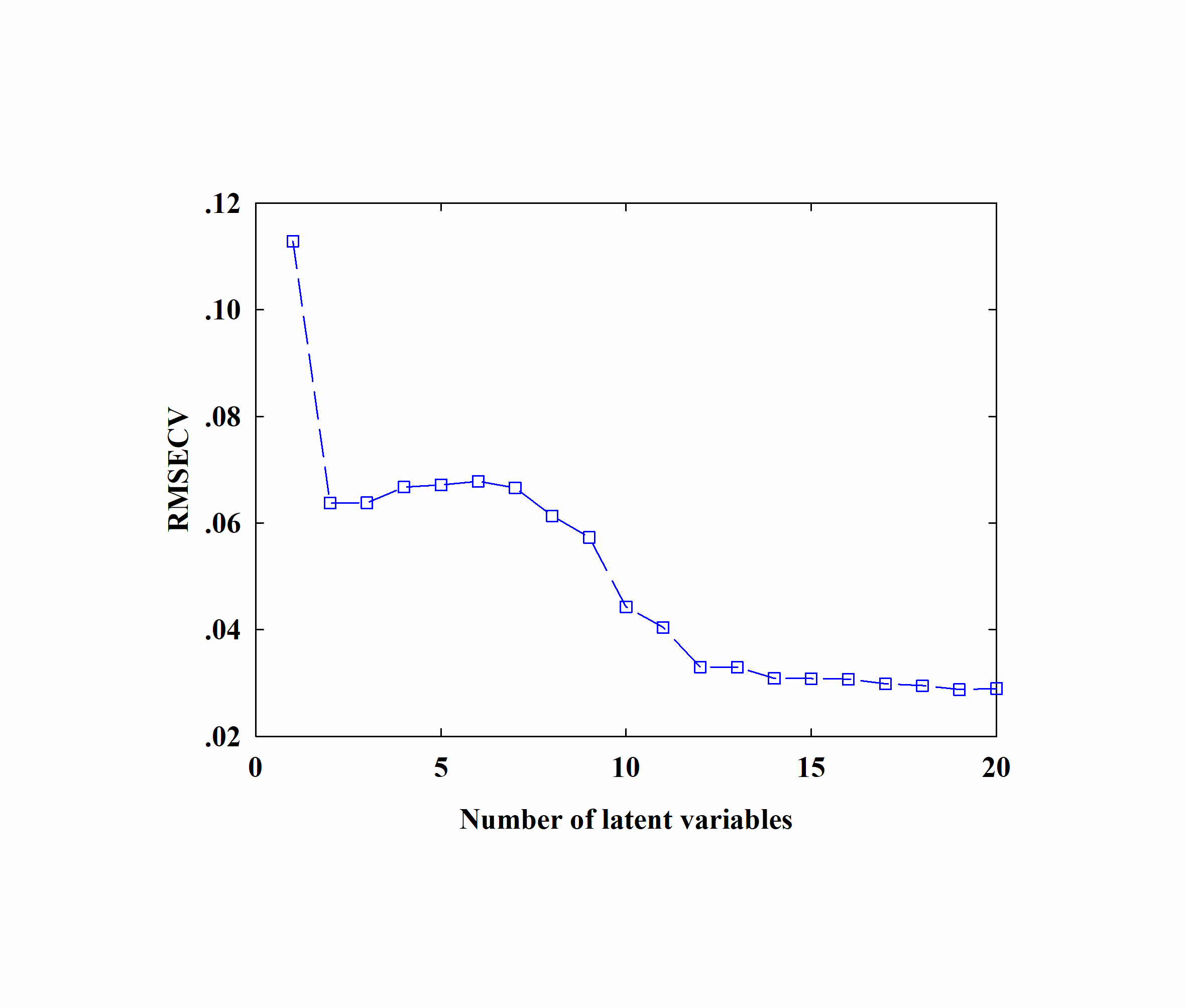
Shengwang Liang: School of Traditional Chinese Medicine, Guangdong Pharmaceutical University, No.280, Higher Education Mega Center, Panyu District, Guangzhou 510006, China; E-mail: [swliang371@126.com](mailto:swliang371@126.com).



**Figure S1.** The effects of number of latent variables in the PLS model on the RMSECV for quantitative analysis of typhaneoside. In order to avoid model over-fitting, the optimal number of latent variables in the PLS model should be determined. The leave-one-out cross validation is used to optimize the number of latent variables. Generally, the RMSECV would decrease as the number of latent variables in the PLS model increases. When the RMSECV reaches the minimum value, the optimal number of latent variables is determined. Abbreviations: PLS, partial least squares; RMSECV, root mean square error of cross validation.

****

**Figure S2.** Correlation diagram of the reference value and predicted value of typhaneoside. The spectra are pretreated by Savitzky-Golay smoothing, and the competitive adaptive reweighted sampling method is used to conduct the variable selection. Thirteen latent variables are used to build the PLS model. The samples in both calibration set and prediction set are predicted by this model. There is a good correlation between the reference value and predicted value. Abbreviations: HPLC, high performance liquid chromatography; NIR, near-infrared; PLS, partial least squares; RMSEC, root mean square error of calibration, RMSEP, root mean square error of prediction.



**Figure S3.** The effects of number of latent variables in the PLS model on the RMSECV for quantitative analysis of isorhamnetin-3-*O*-glucoside. In order to avoid model over-fitting, the optimal number of latent variables in the PLS model should be determined. The leave-one-out cross validation is used to optimize the number of latent variables. Generally, the RMSECV would decrease as the number of latent variables in the PLS model increases. When the RMSECV reaches the minimum value, the optimal number of latent variables is determined. Abbreviations: PLS, partial least squares; RMSECV, root mean square error of cross validation.