Study Protocol (version:2, date:21/01/2022)

Project Title: Diabetes screening and monitoring using tongue images and self-reported symptoms: a machine learning approach

Research protocol number: FRG2/17-18/095； RGC Protocol 12100320 (application submitted)

Principal Investigator

Name: Dr Jason NG

Title: Associate Consultant, Department of Medicine, Queen Elizabeth Hospital

Lead Principal Investigator

Name: Dr ZHANG Shiping

Title: Assistant Professor, School of Chinese Medicine, Hong Kong Baptist University

Study Background

a. Work done by others.

Diabetes mellitus (DM) is a major non-communicable disease. In 2014, the global prevalence of DM was 8.5% in the adult population, and 422 million adults were living with DM (World Health Organization, 2016). In Hong Kong, about 700,000 people or 10% of the population have DM (Quan et al., 2017). Diagnosis and self-management of DM is important. Currently, detection of diabetes requires blood tests, which is costly and inconvenient, especially for elderlies. In view of the fact that a large number of the human population might be affected by diabetes, a convenient and in-expensive method for diabetes screening and monitoring would be highly desirable. Tongue diagnosis has been used in Chinese medicine as a routine diagnostic method, and it has recently been studied for detection of DM and diabetic retinopathy (DR). According to the recent studies by several groups, tongue images obtained through specialized tongue image capture device can detect both DM and DR with an accuracy rate between 80% to 90% (Zhang D. et al., 2017; Zhang et al.,2013; Zhang J. et al., 2017; Li et al., 2016; Que et al., 2016). However, there are certain limitations associated with these studies. For example, there was no description of the subjects with diabetes and those without diabetes, so it is not known if they were matched for such confounding factors as age and other diseases. Moreover, the specialized device is costly and it is not available on the market for layman use. Nevertheless, these studies have provided important indications that tongue images may be used as a convenient tool for detection and monitoring of DM.

Traditional Chinese medicine (TCM) pays great importance to tongue diagnosis, and the tongue can be an important biomarker of DM for several reasons. First, thirst and dry mouth are typical of DM, which can be reflected in the tongue. Moreover, fungal and bacterial infections of the mouth, as well as compromised microcirculations, often occur in DM patients (Jiang et al., 2012; Lu et al., 2016; Hu et al., 2015; Sun et al., 2013; Hsu et al., 2016). All these factors will influence the appearance of the tongue due to changes in hydration, microcirculation and bacterial colonization. Indeed, it has been reported that yellow tongue coating is associated with diabetes mellitus in Japanese non-smoking men and women (Tomooka, 2017). In a study that that analyzed 5930 patients, Tong and his colleagues reported that the majority of DM subjects had thick and/or greasy fur with red or dark red tongue body (He et al., 2013).

HbA1c test, the haemoglobin A1c or glycated haemoglobin test, is an important indication of how well diabetes is being controlled. This test is considered a better means of diagnosing diabetes than measures of glucose levels in the blood. The diagnosis of diabetes is made if the A1c level is ≥6.5%. (The International Expert Committee, 2009). However, the diagnosis of prediabetes is debatable. The Position Statement of the American Diabetes Association defined prediabetes as a fasting plasma glucose between 100 and 125 mg/dl (5.6 - 6.9 mmol/l) and/or a 2-h plasma glucose during oral glucose tolerance test 140 - 199 mg/dl (7.8 - 11.0 mmol), and/or a HbA1c level 5.7 - 6.4%, noting that there is potential discordance between tests (Buysschaert et al., 2016). Clinically, HbA1c reading may be classified into four ranges: <6% (normal), 6-6.4% (prediabetes), 6.5-8.9% (diabetes) and ≥ 9% (diabetes with high HbA1c) (Gaber et al., 2017; Weng, 2017). It has been shown that high level of HbA1c (≥ 9%) is associated with lipid profile of higher risks and actual higher cardiovascular risks (Sherwani et al., 2016; Idrees Khan and Ashfaq, 2018).

So far, there has been no report of using smartphone and tongue image for screening or monitoring DM. There is, however, reports of using smartphone and retinal image for diagnosis of DM and DR (Rajalakshmi et al., 2018). Such technique, however, still requires the use of special device for retinal image collection, limiting its use to professional people only.

To summarize, the current technologies in non-invasive diabetes screening require specialized equipment, and a software solution with the use of smartphone is not yet available. Previous studies suggested that it might be possible to use tongue image for screening diabetes using HbA1c as a marker. However, whether tongue images can help to differentiate different levels of HbA1c, such as pre-diabetes (HbA1c 5.7 - 6.4%) and diabetes with high HbA1c (≥9%), has yet been determined.

b. Work by the investigator

b.1. Study of rater reliabilities of smartphone tongue images

We have developed a method with an Android App for obtaining good quality tongue images using smartphones (Zhang and Parcus, 2017, patent pending). Recently, we catch up with the rapid development of smartphone technologies and simplified this method on smartphones of both Android and iOS systems, without the need for using the App. We validated our new method by measuring its intra-rater reliability between real tongue observation and image observation (Wang et al., submitted). We have found that certain components of tongue features possess higher intra- and inter- rater reliability than the others. Specifically, Intra-rater reliabilities of TCM feature assessments between real tongue and tongue image were good to very good (Cohen's Kappa between 0.7-1.0), except for the color of tongue body and for slippery tongue fur. Inter-rater reliability for tongue images was moderate for tongue coating (Gwet's AC2 between 0.49-0.55), and fair for color and features of tongue body (Gwet's AC2: 0.34). Therefore, certain tongue features can be better retained by smartphone images, while others are more variable. These variations are presumably due to factors of lighting, screen display setting, as well as individual viewer's interpretation, and so forth.

b.2. Tongue diagnosis of abnormal level of HbA1c - proof of concept study

As the first step, we had successfully trained an algorithm that can performed tongue segmentation automatically with 2700 tongue images. Then we trained an algorithm to differentiate subjects with abnormal level of HbA1c (≥6.5%) from subjects with normal HbA1c (<6.5%). The training image set consisted of a total of tongue images from 420 subjects, of which 210 had abnormal HbA1c regarding (≥6.5%) and 210 had normal reading (<6.5%). About 2/3 of the subjects were from a population with high probability of diabetes in endocrine/diabetic clinics, and the remaining 1/3 were from non-specific sources including nursing homes, health check centers, respiratory and nephrology departments. By using transfer learning to classify images with the Resnet101 model, we successfully trained an algorithm that could distinguish between normal and abnormal cases with high accuracy. For a test image set composed of 56 randomly selected images (abnormal HbA1c n=28, mean age=61±1.9; normal HbA1c n=28, age=55±2.7), the results of the algorithm show a sensitivity of 83% and a specificity of 87%, with an overall accuracy of 80.4%. These results demonstrate that tongue images taken by smartphone can indeed be used to detect HbA1c abnormality.

b.3. Tongue diagnosis of abnormal level of HbA1c – feature recognition

To study the image features that contribute to the detection of abnormal HbA1c, heat maps have been generated for image pixels through which machine learning has identified high probability of correlation of normal/abnormal HbA1c level.

Anatomically, the dorsal surface of the tongue is divided by the sulcus terminalis into the anterior two-thirds (oral part), and the posterior one-third (pharyngeal part). TCM tongue diagnosis examines mainly the oral part, which is covered by a connective tissue core with overlying stratified squamous epithelium on the surface. The epithelium on the oral part of the tongue forms three types of papillae, and they have been named after their appearance: filiform, fungiform, and foliate papillae. In TCM, tongue fur refers to the keratinized tip of filiform papillae and dead epithelial cells, and its appearance is also affected by oral bacteria, blood metabolites and salivary secretion from mucous and serous glands (Hu et al., 2015; Washio et al., 2005). The fungiform papillae are highly vascularized, and they provide the based color of the tongue body in TCM diagnosis. The foliate papillae are located at the edge of the tongue posteriorly and have an insignificant contribution to tongue coating. Taken together, we postulated that the appearance of filiform papillae and fungiform papillae, together with dead epithelial cells and the surrounding bacteria, play a significant role in characterizing the tongue images of diabetic subjects.

A close examination of areas corresponding to the heat maps has revealed subtle differences in textural features and their geographical distribution in tongue coating of the two groups. We have found four different types of features likely to be used in machine classification: (i) "thick fur", which is characterized by long and enlarged individual filiform papillae distributed across the entire tongue surface; (ii) greasy fur, which is thick and smooth; (iii) partial or whole mirror-like tongue surface, which is characterized by the shrinkage of filiform papillae; (iv) a combination of two or three of the fore mentioned features. We noticed that the "thick fur" associated with abnormal HbA1c sometimes might be different from the traditional Thick Fur in TCM, in that the long and enlarged filiform papillae might not cover the "body" or the fungiform papillae of the tongue, and the tongue surface appears to be coarse. In this connection, it is worthwhile to note that a previous study had reported over 80% of DM patients had thick and (or) greasy fur of some sort (Tomooka et al., 2017). While one may argue that thick fur and greasy fur are not exclusive to DM or abnormal level of HbA1c, it appears that there are additional subtle differences within the general categories of thick fur and greasy fur, such as the appearance of individual filiform papillae, according to our preliminary data shown above. In this regards, it is interesting to note that we have found tongue image taken with smartphone can reveal more detailed features than visual inspection of patient's tongue during clinical routine practice (Wang et al., submitted).

With the understanding that tongue features alone could only indicate certain probability for a particular range of HbA1c, we had developed a hybrid model for a more accurate prediction of HbA1c range. As TCM believes that any disturbance within the human body must have manifestations outside the body, we combine tongue images with the results from a TCM symptom questionnaire using machine learning techniques. We have filed a patent application for this method (Zhang et al. 2018). The implementation details of such method will be described in the Methodology section.

To summarize, we have developed a method for taking tongue images using smartphone, which can reveal more detailed features than conventional clinical tongue inspection. Using these images, the PI's team had documented the validity and reliability of using smartphone tongue images in place of real tongue in TCM tongue diagnosis. In our proof of concept study, using smartphone tongue images labeled with HbA1c reading, we confirmed previous findings that tongue images could be used for screening HbA1c abnormality, with good sensitivity and specificity. Based on the heat maps of images that had been correctly classified by the trained algorithm, we have begun to identify certain characteristic tongue fur features of HbA1c abnormality presumably used by machine learning.

There are many limitations of the preliminary study. First, the sample size of the present study is small, which limits the external validity of our findings. Second, the characteristics of subject groups have not been well defined, in that their health status and comorbid diseases have not been well documented. Third, the tongue features responsible for the classification of HbA1c have not been well defined. Last but not the least, the pilot study could only differentiate normal versus abnormal HbA1c, whereas prediction of different levels of HbA1c could be more desirable in monitoring the development of DM.

Therefore, it is our plan in this study to address these specific limitations with the following objectives. The results of this study will enable us to develop a practical App for diabetes screening and monitoring.

Study Objective

The aim of the study is to develop an algorithm for diabetes screening, with the following objectives:

(1). To determine the sensitivity and specificity of tongue images taken with smartphone in predicting abnormal HbA1c (≥6.5%);

(2) To determine tongue image features responsible for the classification of normal and abnormal levels of HbA1c (≥6.5%);

(3) To determine the sensitivity and specificity of tongues image in predicting four different levels of HbA1c: <6% (normal), 6-6.4% (prediabetes), 6.5-8.9% (diabetes) and ≥ 9% (diabetes with high HbA1c);

(4) To determine the sensitivity and specificity of combining image analysis results with the results from a TCM symptom questionnaire in predicting the four levels of HbA1c.

Hypothesis

Our working hypothesis is that different tongue coating features may be associated with different stages of diabetes, as indicated by different levels of HbA1c； and different combination of symptoms from a TCM point of view may also be associated with different levels of HbA1c. Thus, combining tongue image with TCM symptoms may allow a machine learning model to build an algorithm for HbA1c prediction with reasonable accuracy.

The design of this protocol is in compliance with ICH-GCP guidelines. The design of these experiments has taken into considerations from previous reviewers, with special consideration on sample size estimation, subject recruitment, subject matching and control of confounding factors.

Inclusion criteria

The inclusion criteria is adult subjects with HbA1c test results from a laboratory that meets the ISO 15189 standard, such as those laboratories used by the Hospital Authority of Hong Kong. The test report must be within a two-week period before or after the collection of tongue images and questionnaire.

Subjects will be recruited from public hospitals in Hong Kong （QE Hand QMH）, Guangdong (Guangdong Provincial TCM Hospital) and HKBU Chinese medicine clinics (Hong Kong Baptist University Mr. & Mrs. Chan Hon Yin Chinese Medicine Specialty Clinic and Good Clinical Practice Centre at Kowloon Tong, and Hong Kong Baptist University – Jockey Club Chinese Medicine Disease Prevention and Health Management Centre at Prince Edward). We have established collaboration with Dr. Ng from Queen Elizabeth Hospital, which is the largest hospital in Hong Kong. The majority of subjects will be recruited from the diabetic clinics. To control for confounding factors such as other chronic diseases (e.g., chronic obstructive pulmonary disease [COPD], renal failure, strokes, etc), we will also recruit subjects from non-diabetic clinics of HKBU mentioned above and from nephrology in-patient departments of Guangdong Provincial TCM Hospital. Subjects must be able to sign the consent form and cooperate in tongue image collection. If subjects do not already have a recent HbA1c test, and they are planning to do the test for health reasons, they will be offered the test free-of-charge. At HKBU out-patient clinics, Chinese medicine practitioners will be informed of our study, and patients who already have or are likely to have a HbA1c test will be invited to join this study by the practitioners. To recruit subjects potentially having a normal HbA1c reading at the HKBU out-patient clinics, free HbA1c tests will be offered to adult subjects who are interested in having HbA1c test as a routine check-up free of charge. The procedures will be performed by a laboratory that meets the ISO 15189 standard, and drawing of blood samples will be performed by a qualified venipuncturist of the laboratory.

Exclusion criteria

Subjects who are unable to give consent, unable to answer the questionnaire or to cooperate in tongue image collection will be excluded. We will not include subject who are unable to understand written Chinese or English.

Tongue image collection

Tongue image collection will be carried out by trained research assistants based on our published protocol (Wang et al., submitted), which consisted of the following steps. Step 1, start the phone's camera function without using any filter or artificial intelligence (A.I.) function; Step 2, turn on the flash; Step 3, aim at the lips of the subject by holding the phone directly in front of, and 15-20 cm away from, the subject with a 45 degree angle; Step 4, tap on the screen to adjust focus when the subject protrudes the tongue and then take a picture; Step 5, upload the image in full resolution to the online questionnaire system (Qualtrics survey software).

Chinese medicine questionnaire collection

A questionnaire in electronic/paper form will be used to collect symptoms relevant to Chinese medicine pattern diagnosis of diabetes. The questions have been selected from previous studies on the most commonly appearing symptoms of diabetes from Chinese medicine perspectives (Hsu et al., 2016; Zhou et al., 2012; Zhao et al., 2017). Age, gender, weight, height , duration of diabetes, family history of diabetes and any comorbid disease, Hb reading (if available) will also be taken into consideration. Recent history of smoking is also recorded as a potential confounding factor for tongue coating features (Tomooka et al., 2017)

Data entry and verification

To ensure accuracy, image, questionnaire answers and HbA1c (and Hb, if available) readings will be uploaded to the server at the same time. A patient identification number for the trial will also be generated to facilitate data verification. Such identification number consists of information on data collection date and outpatient number, etc. However, it will not contain any information related to patient's HKID, full patient name, HN, MRN, DOB, address or phone number.

The data will be collected by an online questionnaire service using Qualtrics and then store at a local computer at SCM/HKBU. A paper version of the questionnaire may be used when necessary, and the filled questionnaire will be kept for 3 years after the study. The collected image will contain no eye features on the photo, and the questionnaire will not include any information for identification of the subject, except a subject ID number for the trial. After completion of the project, data on Qualtrics will be deleted, whereas data on the local computer will be kept for seven years after the study has been published, as per institution requirement.

Study design

This is a cross-sectional design looking at the relationship between tongue image pattern and HbA1c reading. Age, gender, weight, height , duration of diabetes, family history of diabetes and any comorbid disease will be recorded. The level of hemoglobin and blood lipid profile will also be recorded if the information is available. Any acute repertory or digestive illness, as well as smoking habits will also be noted. An electronic questionnaire (using Qualtrics survey software) based on published TCM symptoms of diabetes and the abovementioned information will be used for data collection.

Data processing and analysis

Tongue segmentation

The images containing the tongue and its surrounding area will be processed for segmentation of the tongue area. This segmentation is carried out by a computer algorithm developed in-house by machine learning.

Machine learning

Two approaches will be used in machine learning. In the first approach, we will first perform image classification of either normal or abnormal HbA1c and generate the probabilities for the classification using convolutional neural networks (CNNs) (Anwar et al., 2016; Ødegaard et al., 2016). Then we will try to classify the images into four different classes according to their HbA1c level: <6% (normal), 6-6.4% (prediabetes), 6.5-8.9% (diabetes) and ≥ 9% (diabetes with high HbA1c). The probability data of image classification will be combined with data from the questionnaire as variables using fully connected CNNs layers to determine which one of the four HbA1c levels the subject belongs to (Osia et al., 2018). The advantage of this approach is that it allow us to determine the sensitivity and specificity of the classifications using tongue image alone. However, a drawback of this approach is that it may lose some important features of the tongue in classification with questionnaire data. Alternatively, we will extract features from each tongue image using CNNs, and use one hot encoding to transfer the features into one dimensional matrix, which enable the blending of image data with questionnaire data for regression analysis using neural network (Fig.1; Buckman et al., 2018).

A separate testing data set, which contains at least 200 negative and 200 positive images, will be used to evaluated sensitivity and specificity of the algorithms obtained. In the final algorithm, sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) will be reported.

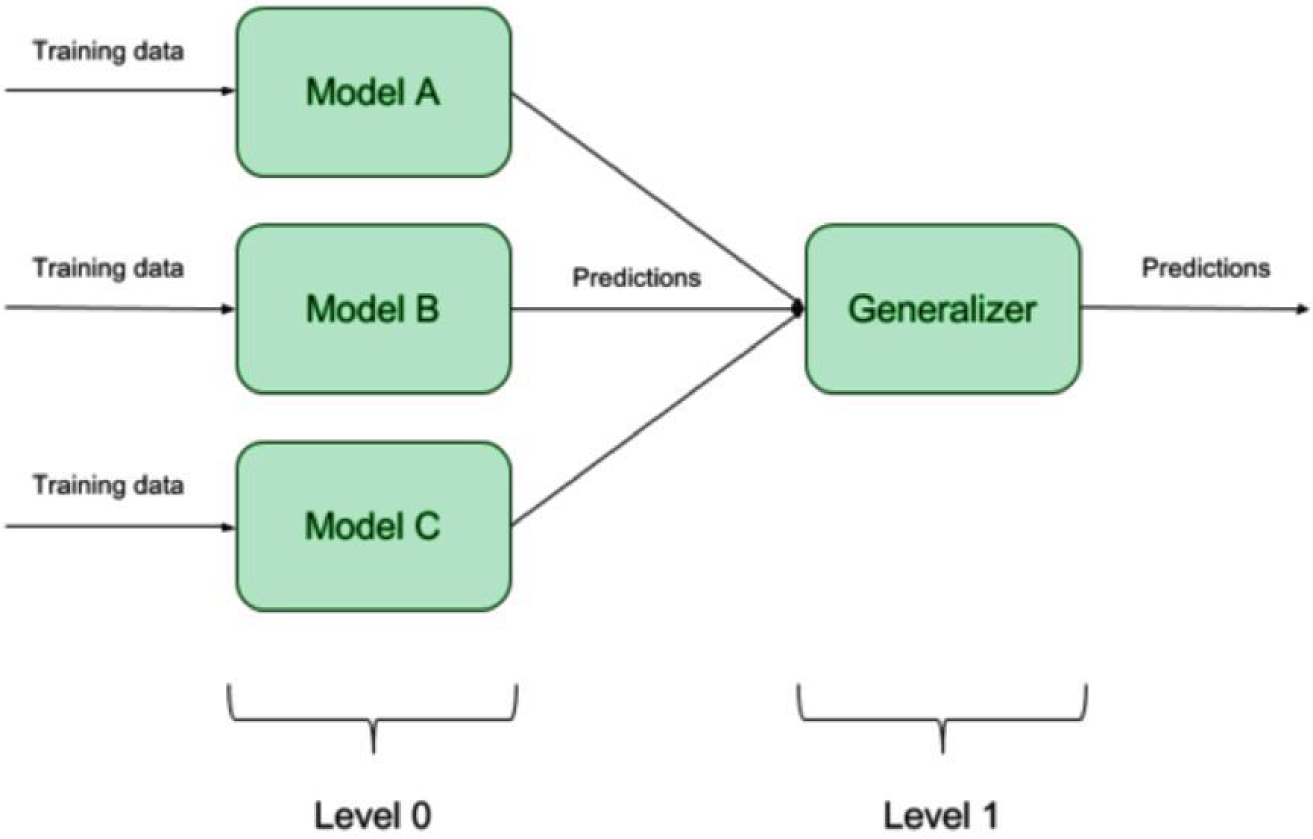

Fig. 1. Schematic diagram showing results from different inputs are blended together to generate a final model through Stacking and Blending of the best models selected.

Potential problems and experimental alternatives

Potential problem 1:

It is highly likely that the tongue image features are not exclusive to abnormal level of HbA1c, and they are also present in smokers and some acute and chronic diseases, such as COPD, upper respiratory infection and chronic kidney or liver diseases.

Experimental alternatives:

We have now included a section for smoking history and known acute or chronic diseases in our questionnaire. If we found a significant correlation between certain confounding variables with abnormal HbA1c tongue diagnosis, we would warn the future user about the unreliability of the algorithm when used in a population with the confounding condition. Furthermore, we would adjust the weighting of parameters in the questionnaire to achieve a prediction model with reasonable specificity and sensitivity, even for populations with the said condition.

Potential problem 2:

Machine classification results from the collected tongue image data set show poor specificity and sensitivity in predicting the four HbA1c levels, even after the inclusion of parameters in TCM questionnaire.

Experimental alternatives:
As we will collect blood lipid profile of subjects from the hospital when available, we can analyze our results as we go. When our total sample population reaches 2,000, and we still see a poor specificity and sensitivity, we will include blood lipid profile test to our community subjects. By adding blood lipid profile as parameters in our regression model, we should be able to enhance the specificity and sensitivity of the prediction, as previous studies have found a close correlation between high level of HbA1c and dyslipidemia (Khan et al., 2007). Of course, we do not wish this to happen because it will lower the value of our prediction model, as it requires an invasive procedure to obtain blood lipid profile. Nevertheless, we are confident that at least a non-invasive model for detection of abnormal level of HbA1c ($\geq$ 6.5%) can be produced, based on the results from our proof-of-concept study.

Potential problem 3:
The image features responsible for the classification of normal and abnormal levels of HbA1c ($\geq$6.5%) could not be readily identified.

Experimental alternatives:
Extraction of visual features that can provide a semantic and robust representation of tongue images with abnormal levels of HbA1c is highly desirable. The greasy and thick coating, which we have identified as features for the classification, needs to be further defined. To this end, we will try different techniques proposed in the literature (Ching et al., 2018), including assigning example-specific importance scores, matching or exaggerating the hidden representation, activation maximization and latent space manipulation, etc.

Primary and Secondary Outcome

Primary Outcome: Tongue image features.
We will extract tongue image features and perform image classification of either normal or abnormal HbA1c and generate the probabilities for the classification using convolutional neural networks (CNNs). Then we will try to classify the images into four different classes according to their HbA1c level: <6% (normal), 6-6.4% (prediabetes), 6.5-8.9% (diabetes) and $\geq$ 9% (diabetes with high HbA1c).

Secondary Outcome: Symptom patterns.
Questionnaire data will be combined with image data for regression analysis

Sample size

It has been suggested that training data size is related to the number of predictor variables, total sample size, and the fraction of positive samples/total sample size. In image classification using deep learning, a rule of thumb is 1,000 images per class (Warden, 2017). In the identification of body parts from CT images, (Cho et al., 2015) showed that 200 images produced an average accuracy of 95.67%, and 1000 images gave an average accuracy of 97.25%; Our study on tongue segmentation using 2700 images found an accuracy over 98% (Wang et al, unpublished observation). Since we have four categories (normal, prediabetes, diabetes with HbA1c between 6.5-8.9%, and diabetes with HbA1c ≥ 9%) of images to differentiate, we have estimated that about 4,000 images will be required. We believe such sample size should give reasonable accuracy as well as good external validity in training the image classification system. Furthermore, we will take two images from each subject when possible, and use image rotation and data simulation techniques to increase the size of the data set (Ødegaard et al., 2016). We plan to recruit 1,500 from QEH, and 1,500 from QMH. Since most subjects in QEH and QMH are expected to be in the abnormal range of HbA1c, we will also recruit 1,000 subjects who are expected to be in the normal range of HbA1c outside the HA system, i.e., from HKBU clinics and the Guangdong Provincial TCM Hospital. Furthermore, although we only study 4,000 subjects, we could simulate over ten thousand images and subject data using data augmentation techniques for machine learning.

The total number of subjects required for HbA1c level prediction can also be estimated based on a multiple regression model, although in practice the prediction may be achieved by machine learning rather than mathematical modeling. Currently available software can only calculate sample size for multiple linear regression. Since our study consists of both continuous and categorical variables, and we do not know whether the final regression model is linear or multi-model in nature, we also use a rule-of-thumb method instead of the software method. Existing guidelines suggest a minimum required sample size based on ratio between number of independent variables and number of cases such as 30 to 1 (Suresh et al., 2012). If we considered the tongue image is a variable as a whole, and the questionnaire contains about 49 variables, the minimum required sample size would be 1,500 (30x50). However, if we considered multiple tongue features as individual variables, the total number of variables would well exceed 50. Nevertheless, the 4,000 subjects proposed should meet the minimum required sample size.

Ethical Concern

The major ethical issues here are patient privacy in relation to tongue image taking. We will ensure that the image only includes the area surrounding the tongue, excluding other unnecessary part of the face. We will also restrict access to the tongue images taken to our own research use.

To protect participants' privacy, all research data would be handled in line with HA / Hospital's policy in handling / storage / destruction of patients' medical records. They would be locked in cabinets where the department or ward keeps patients' confidential information. Electronic data should be saved in secured computer of the hospital with restricted access. The protocol complies with ICH-GCP. All identifiable personal data will be anonymized and will follow the HA policy on handling of patient data privacy.

References:

1. Anwar, S. M., Majid, M., Qayyum, A., Awais, M., Alnowami, M., & Khan, M. K. (2018). Medical image analysis using convolutional neural networks: a review. Journal of medical systems, 42(11), 226.
2. Buckman, J., Roy, A., Raffel, C., & Goodfellow, I. (2018). Thermometer encoding: One hot way to resist adversarial examples.
3. Buysschaert, M., Medina, J. L., Buysschaert, B., & Bergman, M. (2016). Definitions (and Current Controversies) of Diabetes and Prediabetes. Current diabetes reviews, 12(1), 8-13.
4. Ching, T., Himmelstein, D. S., Beaulieu-Jones, B. K., Kalinin, A. A., Do, B. T., Way, G. P., ... & Xie, W. (2018). Opportunities and obstacles for deep learning in biology and medicine. Journal of The Royal Society Interface, 15(141), 20170387.
5. Cho, J., Lee, K., Shin, E., Choy, G., & Do, S. (2015). How much data is needed to train a medical image deep learning system to achieve necessary high accuracy?. *arXiv preprint arXiv:1511.06348*.
6. Hsu, P., Huang, Y., Chiang, J., Chang, H., Liao, P., & Lo, L. (2016). The association between arterial stiffness and tongue manifestations of blood stasis in patients with type 2 diabetes. BMC Complementary And Alternative Medicine, 16(1). http://dx.doi.org/10.1186/s12906-016-1308-5
7. Hsu, P., Huang, Y., Chiang, J., Chang, H., Liao, P., & Lo, L. (2016). The association between arterial stiffness and tongue manifestations of blood stasis in patients with type 2 diabetes. BMC Complementary And Alternative Medicine, 16(1). http://dx.doi.org/10.1186/s12906-016-1308-5
8. Hu, J., Han, S., Chen, Y., & Ji, Z. (2015). Variations of Tongue Coating Microbiota in Patients with Gastric Cancer. Biomed Research International, 2015, 1-7. http://dx.doi.org/10.1155/2015/173729.
9. Idrees Khan MD, Ashfaq F (2018) Variation of Glycosylated Hemoglobin (HbA1C) Cutoff among type-2 diabetic nephropathy Patients. J Endocrinol Diab. 5(6): 1-5. DOI: 10.15226/2374-6890/5/6/001118.
10. International Expert Committee report on the role of the A1C assay in the diagnosis of diabetes. (2009). Diabetes care, 32(7), 1327-1334.
11. Jiang, B., Liang, X., Chen, Y., Ma, T., Liu, L., & Li, J. et al. (2012). Integrating nextgeneration sequencing and traditional tongue diagnosis to determine tongue coating microbiome. Scientific Reports, 2. http://dx.doi.org/10.1038/srep00936.
12. Khan, H. A., Sobki, S. H., & Khan, S. A. (2007). Association between glycaemic control and serum lipids profile in type 2 diabetic patients: HbA 1c predicts dyslipidaemia. Clinical and experimental medicine, 7(1), 24-29.
13. Li, J., Zhang, D., Li, Y., & Wu, J. (2016). Multi-modal Fusion for Diabetes Mellitus and Impaired Glucose Regulation Detection.arXiv preprint arXiv:1604.03443.

14. Lu, H., Ren, Z., Li, A., Zhang, H., Jiang, J., & Xu, S. et al. (2016). Deep sequencing reveals microbiota dysbiosis of tonguecoat in patients with liver carcinoma. Scientific Reports, 6, 33142. http://dx.doi.org/10.1038/srep33142.
15. Ødegaard, N., Knapskog, A. O., Cochin, C., & Louvigne, J. C. (2016, May). Classification of ships using real and simulated data in a convolutional neural network. In 2016 IEEE Radar Conference (RadarConf) (pp. 1-6). IEEE.
16. Osia, S. A., Taheri, A., Shamsabadi, A. S., Katevas, M., Haddadi, H., & Rabiee, H. R. (2018). Deep private-feature extraction. IEEE Transactions on Knowledge and Data Engineering.
17. Parcus R, Chow KL, Zhu HL, Yu Q, Zhang SP. Development of a mobile phone based tongue image acquisition system. In: 2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI). ; 2017:1-6. doi:10.1109/CISP-BMEI.2017.8302299
18. Quan, J., Li, T. K., Pang, H., Choi, C. H., Siu, S. C., Tang, S. Y., ... & Leung, G. M. (2017). Diabetes incidence and prevalence in Hong Kong, China during 2006–2014. Diabetic Medicine, 34(7), 902-908.
19. Rajalakshmi, R., Subashini, R., Anjana, R. M., & Mohan, V. (2018). Automated diabetic retinopathy detection in smartphone-based fundus photography using artificial intelligence. Eye, 32(6), 1138.
20. Sherwani, S. I., Khan, H. A., Ekhzaimy, A., Masood, A., & Sakharkar, M. K. (2016). Significance of HbA1c test in diagnosis and prognosis of diabetic patients. Biomarker insights, 11, BMI-S38440.
21. Sun, Z., Zhao, J., Qian, P., Wang, Y., Zhang, W., & Guo, C. et al. (2013). Metabolic markers and microecological characteristics of tongue coating in patients with chronic gastritis. BMC Complementary And Alternative Medicine, 13(1). http://dx.doi.org/10.1186/1472-6882-13-227.
22. Suresh, K. P., & Chandrashekara, S. (2012). Sample size estimation and power analysis for clinical research studies. Journal of human reproductive sciences, 5(1), 7.
23. Tomooka, K., Saito, I., Furukawa, S., Maruyama, K., Eguchi, E., Iso, H., & Tanigawa, T. (2017). Yellow tongue coating is associated with diabetes mellitus among Japanese non-smoking men and women: the toon health study. Journal of epidemiology, JE20160169
24. Warden, P. (2017). How many images do you need to train a neural network.
25. Washio J, Sato T, Koseki T, Takahashi N. Hydrogen sulfide-producing bacteria in tongue biofilm and their relationship with oral malodour. J Med Microbiol. 2005;54(Pt 9):889-895. doi:10.1099/jmm.0.46118-0. PMID: 16091443
26. World Health Organization. (2016). Global report on diabetes. World Health Organization.
27. Zhang, B., Kumar, B. V., & Zhang, D. (2014). Detecting diabetes mellitus and nonproliferative diabetic retinopathy using tongue color, texture, and geometry features. IEEE transactions on biomedical engineering, 61(2), 491-501.
28. Zhang, D., Zhang, H., & Zhang, B. (2017). Tongue Image Analysis (pp. 1-335). Springer.
29. Zhang, J., Xu, J., Hu, X., Chen, Q., Tu, L., Huang, J., & Cui, J. (2017). Diagnostic method of diabetes based on support vector machine and tongue images. BioMed research international,
30. Zhang, S.P. and Parcus R. (2017) Method, device and system for obtaining tongue image. Chinese patent application No. CN 106859595A.

31. Kan, H. X., Zhang, L. Y., & Dong, C. W. (2016). A tongue image recognition method for TCM syndrome differentiation of type 2 diabetes. Chinese Journal of Biomedical Engineering, 35(6), 658-664. (In Chinese)

32. Cluster analysis of TCM syndromes in type 2 diabetes. Chinese Journal of Traditional Chinese Medicine, 27(12), 3121-3124. (In Chinese)

33. Zhao, Y. Q., Li, Q. S., Xiang, M. H., Zhang, W., & Zhang, X. R. (2017). Distribution pattern and symptom correlation of TCM syndromes in diabetic retinopathy. Chinese Journal of Traditional Chinese Medicine, 42(14), 2796-2801. (In Chinese)

34. Zhou, D. Y., Zhao, J. X., Mou, X., Liu, W. H., Zhou, D. Y., Liu, Y. H., ... & Wang, S. R. (2012). Cluster analysis of TCM syndromes in type 2 diabetes based on "Zheng". Chinese Journal of Traditional Chinese Medicine, 27(12), 3121-3124. (In Chinese)