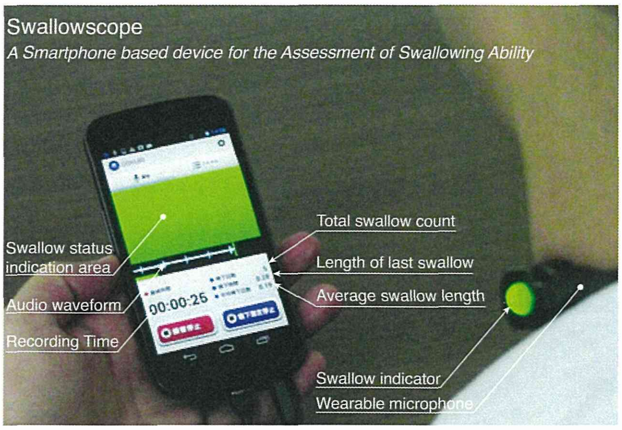

Fig. 4. Overview of the Swallowscope, showing the indicator side of the neck-worn microphone and the display of the smartphone application

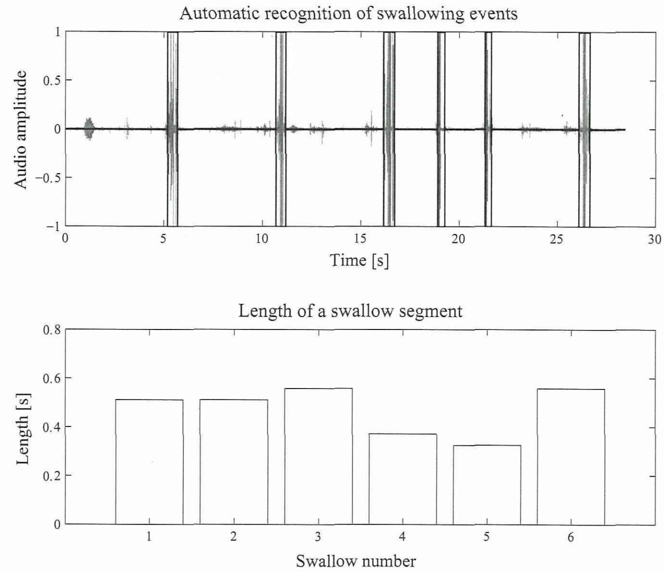

Fig. 5. Real-time recognition of swallowing activities during RSST

In order to find necessary parameters for the swallow activity detection algorithm, we performed RSST on 15 subjects; and to analyse the data for feature extraction, we asked a speech therapist and an experience swallowing sound coder to tag the starting and end points of each swallow during RSST. From the data, we noticed significant variations in lengths of swallow segments from person to person (p-value<0.001), but less variations for the same person. Fig. 3 shows the variation of average swallow lengths of the 15 subjects. The test yielded an average swallowing time of 0.51 s (S.D = 0.16).

B. *Real-time processing of swallowing sound*

The swallowing sound processing consists of 3 phases: preliminary detection of swallow sound pattern, elimination of cough and vocalizations and final detection. In the smartphone based algorithm, audio is captured as 16-bit pcm (Pulse-code modulation) at a rate of 11025 Hz and processed after acquiring into a 512-size buffer.

In the preliminary detection, we capture the sound profiles that are bounded by silent segments longer than 139 ms (3 buffers) to a sound segment. Here, an acquired audio data buffer is considered silent if the maximum value is less than a predefined threshold. As it can be seen from Fig. 2, typically the audio produced at different anatomical levels are separated by silent periods that are shorter than 3-buffers. As a result, a silent period within an audio segment indicates a switch in the anatomical region. After selecting the preliminary audio segment, a continuous wavelet transformation is performed and the amplitudes of the scale 19 is compared against a predefined threshold to detect cough and sections of vocalization. Finally, the audio segment is tested for: (a) total length, which should be between 232 ms (5-buffers) and 700 ms (15-buffers), and (b) number of sections (detected anatomical regions), which should be between 2 and 4. If all the above conditions are satisfied, the audio segment is considered a proper swallow.

Fig. 4 shows the smartphone-based Swallowscope. The swallowing sound is captured from the neck-worn wired contact microphone and the real-time processing outcome is displayed on the screen. The microphone was placed approximately around the c5 level of the cervical spine; however, since the features used for swallow recognition were less sensitive to the microphone location [9], it was often worn by the subjects, themselves. The display statistics consists of the number of swallows recognized, the total time taken for the last recognized sallow and the average time taken for a swallow. The recognition outcome is shown both at the indication area of the smartphone and with the indicating full-colour-LED of the neck-worn microphone with 3 different colours: blue for no swallow or stand-by, green for proper swallow and red for cough.

III. EVALUATION AND RESULTS

In order to evaluate the accuracy of automatic recognition, we used the Swallowscope to evaluate RSST activity of 8 subjects while simultaneously performing VF, whereas the VF images were used as the ground-truth to calculate recognition accuracy. Fig. 5 shows the results of one of the RSST, reproduced with the exact offline version of the real-time algorithm. In this instance, the Swallowscope recognized all the swallowing events accurately.

Fig. 6 shows the precision (fraction of retrieved instances that are relevant) and recall (fraction of relevant instances that are retrieved) of automatic recognition estimated according to the ground-truth obtained from VF evidence. Overall, the automatic swallow recognition algorithm achieved a precision of 83.7% and a recall of 93.9%.

Due to the quantitative nature of evaluation, it is possible to use the Swallowscope to obtain more information about the swallowing process, and Fig. 7 compares the timing characteristics of swallows during RSST of a healthy subject (male, 35) and a dysphagic subject (female, 60's). The healthy subject managed 15 number of swallows within the period of 30 s whereas the dysphagic subject could manage 2 swallows only.

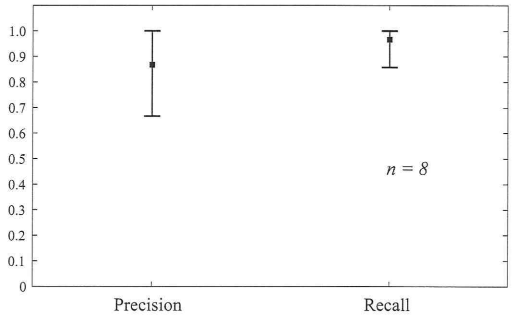

Fig. 6. VFSS evidence-based accuracy for the automatic detection of swallowing activities during RSST

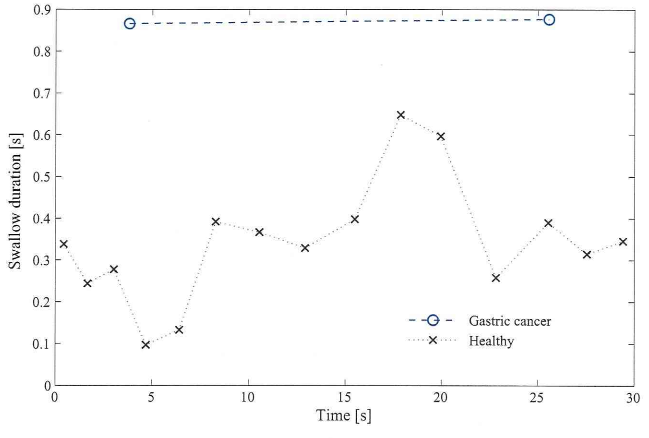

Fig. 7. Comparison of RSST timing characteristics of a healthy subject and a dysphagic subject due to gastric carcinoma

IV. DISCUSSION

Swallowscope is the smartphone based device we developed to continuously monitor swallowing activities and assess the swallowing ability. In the current stage where we focused more on the RSST, we wanted to increase the recall: the true positive rate, to make sure the Swallowscope's categorization of not risky could be accepted with high confidence. The standard screening techniques used to assess swallowing ability have 2 main limitations: the person who conducts the screening test needs to be an expert to identify various outcomes of the test including correctly recognizing swallows and the outcome of the tests are subjected to qualitative interpretation of the examiner. The Swallowscope on the other hand does not require any special knowledge nor training and the results it produce are quantitative. As a result, they are reproducible and comparable. Our current algorithm for instance is capable of recognizing precise temporal characteristics of swallowing activities. It can accurately determine when did the swallow has occurred and how long the swallow was. As a result, it could be possible to define more definitive parameters to evaluate swallowing ability. As seen in Fig 5, we could analyse the progress of swallowing activities, whether the subject is slowing down by using gap between swallows, and whether the time for a swallow is increasing from the swallow length information. These parameters could indicate whether the swallowing function of the subject is getting unusually tired. As it can be seen from Fig. 7, healthy subject, while managing 15 swallows, did not have any problem maintaining the swallowing rate. And the swallows too took much shorter time to complete, indicating much stronger and quicker swallows. Just as coughing indicates possible aspiration, ability to cough properly is important to make sure that the person can cough any aspirated food particles out. Healthcare workers often would like to know it before feeding someone who is in a risk of aspirating, and we could use the cough detection to determine whether their coughs are strong enough to serve this purpose.

V. CONCLUSIONS AND FUTUREWORKS

We developed a smartphone-based device that can analyse swallowing sounds in real-time and generate quantitative results to assist the bedside screening for swallowing ability. We achieved a very high value for recall, indicating the Swallowscope's categorization of not risky could be accepted with confidence. We further demonstrated the ability to generate quantitative measures about the swallowing ability, and this could help to produce more reliable screening methods.

We are currently working on compiling a large database of healthy and dysphagic swallowing patterns and plan to improve pattern-recognition algorithm to increase the screening accuracy and robustness as well as to directly detect dysphagic swallowing patterns. We also believe it is possible to describe swallowing pressure from the swallowing sound and currently conducting a study that combines swallowing sound, VFSS and high-resolution manometry (HRM).

REFERENCES

[1] E. T. Cunningham Jr and B. Jones, *Normal and Abnormal Swallowing: Imaging in Diagnosis and Therapy*. Springer, 2003, ch. 2.

[2] "Vital statistics in japan-the latest trends," Ministry of Health, Labour and Welfare, Tech. Rep., 2011. [Online]. Available: http://www.mhlw.go.jp/english/database/report.html

[3] O. Ekberg, S. Hamdy, V. Woisard, A. Wuttge-Hannig, and P. Ortega, "Social and psychological burden of dysphagia: Its impact on diagnosis and treatment," *Dysphagia*, vol. 17, no. 2, pp. 139–146, 2002.

[4] P. E. Marik and D. Kaplan, "Aspiration pneumonia and dysphagia in the elderly," *CHEST Journal*, vol. 124, no. 1, pp. 328–336, 2003.

[5] V. Tentzeris, B. Lake, T. Cherian, J. Milligan, and A. Sigurdsson, "Poor awareness of symptoms of oesophageal cancer," *Interactive CardioVascular and Thoracic Surgery*, vol. 12, no. 1, pp. 32–34, 2011.

[6] B. Martin-Harris and B. Jones, "The Videofluorographic Swallowing Study," *Physical Medicine and Rehabilitation Clinics of North America*, vol. 19, no. 4, pp. 769–785, 2008.

[7] S. B. Leder, D. M. Suiter, H. L. Warner, L. M. Acton, and B. A. Swainson, "Success of recommending oral diets in acute stroke patients based on passing a 90-cc water swallow challenge protocol," *Topics in Stroke Rehabilitation*, vol. 19, no. 1, pp. 40–44, 2012.

[8] S. Horiguchi and Y. Suzuki, "Screening tests in evaluating swallowing function," *JMAJ*, vol. 54, no. 1, pp. 31–34, 2011.

[9] M. Nagae and K. Suzuki, "A neck mounted interface for sensing the swallowing activity based on swallowing sound," in *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, 2011, pp. 5224–5227.

[10] S. Morinière, M. Boiron, D. Alison, P. Makris, and P. Beutter, "Origin of the sound components during pharyngeal swallowing in normal subjects," *Dysphagia*, vol. 23, no. 3, pp. 267–273, 2008.

[11] K. Oguchi, E. Saitoh, M. Mizono, M. Baba, M. Okui, and M. Suzuki, "The Repetitive Saliva Swallowing Test RSST as a Screening Test of Functional Dysphagia. (1). Normal Values of RSST," *The Japanese Journal of Rehabilitation Medicine*, pp. 375–382, 2000, In Japanese.