Data Quality in mHealth

Walter Karlen, *Senior Member, IEEE*

***Abstract*—mHealth enables remote health monitoring and facilitates communication and data exchange between health providers and patients, but comes at the cost of reduced data quality. Loss of quality can be multifaceted and includes an increase in missing data and a higher level of uncertainty in data compared to clinical data acquisition. This short review discusses emerging strategies to cope with challenges in keeping mHealth systems that are used by lay users reliable.**

I. Introduction

Home-based health care combined with mobile communication technologies (mHealth) is a solution to bring health care closer to the patient. Across the globe, people are rapidly acquiring low-cost mobile phones as daily tools for communication. Thus, embedding bio-medical diagnostic applications on mobile devices could be easily adopted and distributed. This decentralization of medical assessment and diagnosis is expected to provide relief to the overloaded health care systems. However, mHealth data is highly susceptible to noise and misinterpretation, and could lead to the opposite effect. It is therefore important to ensure quality with proper data management early in the design process of mHealth applications.

II. Data Quality Assurance

For most applications data quality is required to be tested on-line, in an automated way at the point-of-care. Many current systems require the data to be analyzed offline. Often this is performed manually by an expert observer after transmission to a central server. This approach is labor intensive and costly. It does not scale up easily and causes high data loss as is the case of poor signal quality or missing data. Recordings must be excluded as the patients are often not available for reassessment.

The quality of periodic vital sign time series data can be evaluated by means of a signal quality index (SQI) that quantifies the reliability of a signal. It often represents a combination of measures such as normality (how similar the signal is to a standardized template) and regularity (how similar it is to itself over time). For example, electrocardiograms recorded with an mHealth sensor suite can be assessed using an SQI to reduce false alarms [1], and artifacts in pulse oximetry waveforms can be detected by using autocorrelation [2]. The difficulty of defining a SQI is the requirement for a gold standard quality assessment; often only available through multiple expert evaluations and finding an agreement between the raters.

*Research supported by the Swiss National Science Foundation (SNSF).

W. Karlen is with the Institute of Robotics and Intelligent Systems, Department of Health Sciences and Technology, ETH Zurich, Technoparkstrasse 1, 8005 Zurich, Switzerland and can be reached at *walter.karlen@ieee.org*.

Manually entered form data can also be suspect to error or manipulation. Supervised and unsupervised machine-learning techniques and statistical representation can identify abnormal data entry in mHealth survey forms [3]. This work has been extended to use behavioral data [4].

User engagement and feedback is a valuable approach for communicating SQI and other quality measures and indirectly improving data recorded with mobile phones. We have implemented an algorithm to detect the correct finger placement on the phone camera for measurement of reflective pulse oximetry [5]. By providing real-time feedback to the user, recordings with insufficient quality were eliminated at the time of recording. Similarly, by monitoring the variability of tapping times while assessing respiratory rate with a mobile phone, instantaneous feedback on measurement progress can be provided, eliminating the need for repeated measures [6]. This same app also provided animations and haptic feedback to allow better evaluation of the obtained results and compare them in real-time with the patient's respiratory effort.

Conclusion

Non-medical experts are the main users of mHealth. This lack of expertise requires compensation with technological innovation to maintain the integrity of data collected and mHealth systems themselves. Despite the many advances in quality assessment techniques, further research is needed for developing and evaluating new approaches for automated management of system integrity in remote health care settings.

References

[1] G. D. Clifford, C. Arteta, T. Zhu, M. Pimentel, M. Santos, J. Domingos, M. Maraci, J. Behar, and J. Oster, "A scalable mHealth system for noncommunicable disease management," in *IEEE Global Humanitarian Technology Conference (GHTC 2014)*, 2014, vol. 2025, pp. 41–8.

[2] W. Karlen, K. Kobayashi, J. M. Ansermino, and G. A. Dumont, "Photoplethysmogram signal quality estimation using repeated Gaussian filters and cross-correlation.," *Physiological measurement*, vol. 33, no. 10, pp. 1617–29, Sep. 2012.

[3] B. Birnbaum, B. DeRenzi, A. D. Flaxman, and N. Lesh, "Automated quality control for mobile data collection," in *Proceedings of the - ACM DEV '12*, 2012, p. 1.

[4] B. Birnbaum, G. Borriello, A. D. Flaxman, B. DeRenzi, and A. R. Karlin, "Using behavioral data to identify interviewer fabrication in surveys," *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*, p. 2911, 2013.

[5] W. Karlen, J. Lim, J. M. Ansermino, G. A. Dumont, and C. Scheffer, "Recognition of correct finger placement for photoplethysmographic imaging," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society.*, 2013, pp. 7480–3.

[6] W. Karlen, H. Gan, M. Chiu, D. Dunsmuir, G. Zhou, G. A. Dumont, and J. M. Ansermino, "Improving the accuracy and efficiency of respiratory rate measurements in children using mobile devices.," *PLoS ONE*, vol. 9, no. 6, p. e99266, Jan. 2014.