Four datasets and Microsoft Azure

Nazrul Anuar Nayan

I. Objective

1) To update the status of the remaining four patient datasets.
2) To update the learning process of Microsoft Azure and Python.

II. Completed tasks

A. Data structure

Previously, analysis on Fantasia, CapnoBase, MIMIC II, Dialysis 1, Dialysis 2 and Dialysis 3 have been reported. The analysis on CALMS2, PICRAM, Listen and Vortal have not started as the datasets are not available. In this report, the status of these four patient datasets, which are not publicly available, will be updated. Table I, which was last shown in ProgReport4 (1 April 2015) is being updated.

TABLE I
DATASETS UPDATE

	CALMS2	PICRAM	Listen	Vortal
Data updated	2014	2015	2014	2014
Formatted/ ⇑ bspprojects9	✗	✗	✗	✗
Raw data	bspprojects5	bspprojects12	✗	✗
ECG Signal	✓ (only 25)	ICU ✗, ward ✓	✓	✓
PPG Signal	✓	ICU ✗, ward ✓	✓	✓
Accelerometry Signal	✗	✗	✗	✗
Subject Population	post-surgery	ICU-discharged	post-cardiac surg	healthy(young/elderly)
Clinical Setting	post-op ward	ICU, ward	ICU, ward	laboratory
Breathing (spontaneous/ ventilated)	s	ICU s,v / ward s,v	s,v	s,v
No. of records	~ 250	ward (Ox 199, Rb 68)	196	42 young, 14 elderly
Neonates (<1 y.o.)	✗	✗	✗	✗
Paediatrics (1-18)	✗	✗	✗	✗
Young Adult (19-40)	✓ (very few)	✓	✓	✓
Adult	✓	✓	✓	✗
Elderly (>70 y.o.)	✓	✓	✓	✓
Unwell?	✓	✓	✓	✗
Chronically Unwell?	✗	✗	✗	✗
Acutely Unwell?	✓	✓	✓	✗
Ambulatory?	ward ✓	✓, ✗	✓, ✗	
Recording Time	1:30 h	ward 1-2 d	2d ICU, 5d ward	[1]
ECG/PPG Acq. Equip.		PPG: Nonin 4100	ICU (bedside), ward (telemetry)	Philips (ECG,PPG 125 Hz), Nonin (PPG 75 Hz)
Continous RR Signal	IP	✗	ICU - IP, ward ✗	IP
Other Reference RR	nurse	nurse (track & trigger)	ICU (ventilator or manual), ward (nurse)	✗
Avalability	available soon	available	> 2 months	>2 months

1) CALMS2: As shown in Table I, raw CALMS2 dataset are located in 'bspprojects5'. At this moment, I can access the folder cancer_hospital, but unable to view the content of the subfolders. A request has been sent to IT support staff on this matter. Based on the discussion with Mr. Pimentel yesterday, there are 407 patients in CALMS2 but only 80% has waveform data.

2) PICRAM: I can now access bspprojects12, 'ward data' where there are 199 Oxford and 68 Reading patient folder which contain ECG and PPG waveforms. The work to put into them into a standard data format and analysis will begin next week. Mr. Pimentel also explained the reference respiratory rate which are taken using track & trigger every 4 hours. No 'gold standard' respiratory signal in PICRAM datasets.

3) Listen: Mr Charlton and Dr. Bonnici is currently extracting data from the dataset in order to evaluate respiratory rate algorithms. It will take approximately one month for the signal pre-processing. Their next step is to write a publication. These data would be available after the publication. I have put the data– to be available after 2 months in the Table [**?**].

4) Vortal: Mr. Charlton and Dr. Bonnici are finalising a draft of publication analysing Vortal dataset, which will be submitted in 2 weeks. Thus, Vortal dataset will be available after this publication. From my estimation, the data will only be available in 2 months.

B. Microsoft Azure and Python

A discussion has been carried out with Mr. Pimentel on the question whether it is good to upload our version of CapnoBase and MIMIC II, with the algorithms now in the respiratory rate toolbox into cloud. He accepted this parallel computing idea and see the benefits of doing it such as to speed up the developments. He will update us once he has fully read and understood Microsoft Azure. He also mentioned that they might be a way to run Matlab in the system. As for python, I have it installed into my office PC and watched a few online tutorials.

III. Conclusion

This week, the access to PICRAM datasets has been provided. Next week, CALMS2 and PICRAM analysis will start. Further understanding of Microsoft Azure and Python will also be carried out.

Acknowledgment

The author would like to thank Mr. Pimentel and Mr. Charlton for their time and advice this week.