The Use of Expert Annotations to Assess Signal Quality of Phonocardiograms using mStethoscope: a feasibility study


**Abstract**—This paper evaluates the feasibility of using expert annotations to assess signal quality of phonocardiograms captured using a mobile stethoscope (or m-stethoscope) by evaluating the variability in expert annotations of signal quality. Signal quality is typically assessed using quantitative measures that aim to represent signal-to-noise ratio or signal entropy. However, in this paper we investigate the feasibility of learning from experts. Phonocardiograms from 150 volunteers were captured by a basically trained health worker in a controlled environment using a 3M Littmann Electronic Stethoscope 3200 and an iPhone 3G. Two consultant cardiologists annotated the recordings by rating each record on a five-point scale, with 1 corresponding to “Excellent” and 5 corresponding to “Noise”. Cohen’s *Kappa coefficient* is used to assess the intra- and inter-rater agreement. The agreement between annotators was found to be 0.58 and 0.27, for the iPhone and Littmann, respectively.

I. BACKGROUND

Mobile health has the potential to transform healthcare in the developing world [1]. Mobile phones offer an attractive method for remotely monitoring heart conditions through digital phonocardiograms (PCGs). The built-in microphone, recording and transmission capabilities on all mobile phones may allow remote experts to analyse PCG recordings. Further, due to their ease of use, mobile phones could be used by community healthcare workers (CHW) to record heart sounds in the field. It is important in this use-case to provide automated real-time feedback on signal quality to ensure the only recordings of high-quality are relayed for further expert analysis. This study evaluates the feasibility of using expert annotations to assess signal quality.

II. METHOD

This study was approved by the Human Research Ethics Committee based at the Health Science Faculty of the University of Cape Town (HREC REF: 568/2010). It was conducted in the Cardiology Clinic at Groote Schuur Hospital in Cape Town, South Africa. The cardiac auscultation recordings, each one minute in length, were made using a 3M Littmann Electronic Stethoscope 3200 and an iPhone *mStethoscope* [3]. Cardioscan (Zargis) was used to export the cardiac auscultations from the Littmann into uncompressed ‘Wave’ format with a sampling frequency of 4 kHz and an amplitude resolution of 16 bits. Two consultant cardiologists annotated the recordings on a 3-point scale, with 1 corresponding to “Excellent”, 2 corresponding to “Acceptable” and 3 corresponding to “Unfit for diagnosis or further analysis”. Fifty recordings were duplicated in the dataset so as to enable intra-analyst agreement.

III. RESULTS

Cohen’s *Kappa Coefficient* ($\kappa$) is a measure of agreement between annotators for qualitative items [4], defined as:

$$\kappa = \frac{Pr(a) - Pr(e)}{1 - Pr(e)}$$

where $Pr(a)$ is the relatively observed agreement among annotators and $Pr(e)$ is the hypothetical probability of chance agreement. Table 1 shows intra- and inter-annotator agreement for the Littmann and iPhone *mStethoscope*, respectively. The annotators agreed on 144 and 101 iPhone and Littmann recordings, respectively. The mean and standard deviation of the agreed annotations were 2.2±1.3 and 1.2±0.6 for iPhone and Littmann, respectively.

<table>
<thead>
<tr>
<th><strong>TABLE I. COHEN’S KAPPA COEFFICIENT</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Analyst</td>
</tr>
<tr>
<td>A</td>
</tr>
<tr>
<td>B</td>
</tr>
<tr>
<td>A + B</td>
</tr>
</tbody>
</table>

IV. DISCUSSION

This work highlights the difficulty of obtaining a high-quality PCG, especially using the *mStethoscope*, and hence the importance of robust signal quality assessment at point-of-care. The low intra-annotator variability for the Littmann recordings highlights the difficulty of assessing quality in high fidelity cardiac auscultations. Further annotators may improve the inter-annotator agreement.

REFERENCES

