Introduction

This Annex lists the data-fields available from the Airwave Study Tissue Bank electrocardiogram ECG export. The reader should consult the main data-dictionary for a description of the screening protocol itself, explanations of formatting conventions, missing values, and other technical information.

We carried out ECG analyses of most participants at baseline and some participants during the follow-up. The choice of whether an ECG was carried out or not was based on practical and budgetary constraints, not medical criteria.

The ECG was collected on a machine recommended by Professor Peter McFarlane's CARE (Computer Assisted Reporting of Electrocardiograms) team at Glasgow University. We uploaded the data to CARE daily and they returned an interpretation to us, usually within a month. We are grateful to Professor McFarlane and his team for their long-standing contribution to the Study.

The interpretation came in two parts: a summary statement, aimed at the advising the participant whether there is anything in the results that merits further clinical investigation; and a more detailed dataset intended for use by researchers.

The Airwave team are not experts in this subject area and will be especially interested in commentary from specialists to help improve or clarify areas of uncertainty.

Document Configuration

Title	Data Dictionary, Annex C
Subject	Electrocardiogram (ECG) Export
Version	2.0
Status	Draft
Author	Heard, Andy H, Database Manager at Imperial College London.
Published	17/05/2022
Filename	annex-c.docx

Changes since Previous Version

We have relocated the summary interpretation previously in the screening extract into the ECG Detail table.

We have extracted the Minnesota Codes and Group Codes into separate tables.

Additional metrics are available for recently interpreted ECGs.

ECG Detailed Results

The researchers' results are divided into three sets of values:

- Values describing the trace geometry.
- Up to 20 Group Codes, which provide structured commentary on the ECG and which appear to be based on a system proprietary to Glasgow CARE.
- Up to 24 Minnesota Codes, which we understand to be a more widely used classification for ECGs.

Glasgow Group Codes

Each Group Code is represented as three values: a group number (2 digits), a statement type (2 digits) and statement code (5 digits). Taken together, they form a compound key within a table of narrative interpretations. We present here the narrative interpretation only, though raw group-codes are available if required (contact the Airwave team for details).

The statements are intended to be read in order, and the succeeding statement can sometimes refer to the preceding statement. For example, PROBABLE NORMAL VARIANT may appear more than once, in each case it refers to the previous statement.

Minnesota Codes

Each Minnesota Code consists of a three digit number that corresponds to a narrative interpretation. As we have no easy access to a lookup table, we have exported the codes we received from CARE in the format given to us. The hyphens were removed by CARE, though we understand the normal convention is to keep them in.

Missing Values

In other datasets we use a "Contingency" value to explain why certain values are missing (ex-protocol, not-found, not collected etc.). For numeric variables, the contingency codes are negative numbers because valid laboratory assay results, for example, are always non-negative. Within the ECG dataset, however, values can be legitimately negative, and so our usual convention is unavailable.

That said, there is rarely any great story behind the reason for an ECG-related value being missing. It will usually be because the value is not applicable in the circumstances, or there was a technical failure. In the rare circumstances where we do need to distinguish missing values from the results of a technical error, the tables explain the substitution values used.

Interpretation Systems

The interpretation systems used by CARE evolved over the course of time. The first version was used from the beginning (2004) until June 2011; a later version from June 2009 until May 2015 (note the overlap, which resulted from compatibility issues with older ECG carts), and the most recent version from November 2015 onwards.

Later systems included more data and so we have provided a label that states which system was used for each ECG.

A small number of results were interpreted manually from the paper trace. This happened when data being transmitted to CARE was lost because of technical faults. Very limited interpretation of these ECGs was possible.

Results Not Exported at this Release

Two variables that have not been exported are:

- **ecg_cart**: available for 16% of results, this is an identifier for the machine used. Any researcher
- **ecg_comment**: available for 3% of results, this is free format text relating that tends to relate either to the ECG itself ("lead 1 won't stick to patient") or participants' self-reported cardiac diagnoses. Because of the risk of disclosing personal data, we have not exported these results.

If you want access to these data, please get in touch with us.

Trace Geometry

Label	Type	Commentary
ecg_id	NUMBER (17)	This is a unique ID generated by the ECG management system. It appears to be generated from the recording date, time and recording device identity. For manually interpreted results we have constructed **ecg_id** using prefix "MAN" and a coded value of the recording time.
barcode	NUMBER (5)	Health-screening identifier.
when_recorded	DATE	A timestamp from the ECG machine.
interpretation_system	STRING(1)	System version of the interpretation system used. See explanation above.
summary_interpretation	STRING	The summary interpretation. We were advised by Professor McFarlane that only summary reports including the word "ABNORMAL" suggest that further clinical investigation is merited.
lvh_interpretation	STRING	A statement relating to any ventricular hypertrophy diagnosed. This is only available when interpretation = 3; otherwise "EX-PROTOCOL".
minnesota_codes	NUMBER	Quantity of Minnesota Codes reported for this ECG
group_codes	NUMBER	Quantity of Group Codes reported for this ECG
heart_rate	NUMBER	Pulse rate in beats per second.
p_axis	NUMBER	P axis.
qrs_axis	NUMBER	QRS axis.
t_axis	NUMBER	T axis.
qrs_duration	NUMBER	QRS duration.
pr_interval	NUMBER	PR interval.
qt_interval	NUMBER	QT interval.
qtc_interval	NUMBER	QT interval.

LVMI and Voltage Based Indices

These analytic results are available in the trace-geometry table when **interpretation_system** = 2 or 3. The following description of these derived values was provided to us by Glasgow CARE in August 2011.

LVMI Rautaharju

Described in [Rautaharju PM, et al; Am. J. Card. 1988; 62:59-66]. The units are g/m^2.

Males: -36.4 + 0.01 RV5 + 0.02 SV1 + 0.028SIII + 0.182 (T-) V6 - 0.148 (T+) aVR + 1.049 QRSd

Females: $88.5 + 0.018RV5 + 0.053\ SV5 - 0.112\ SI + 0.108\ (T+)\ V1 + 1.7\ (T-)\ aVF - 0.094\ (T+)\ V6$

Where Q, R, S voltages are all positive (μV) and S=max(Q,S) and R=max(R,R').

The equation for black females is different from the one above, and has not been used in this extract.

LVMI F. Huwez

Described in [Huwez F; PhD thesis, University of Glasgow, 1990]. The units are g/m^2.

Males: $99.6615 + 0.03358\ SV5 + 0.10121\ (T+)V1 - 0.4231\ (T-)V1 + 0.12692\ (T-)V6$.

Females: $115.1371 - 0.04284\ RIII - 0.02134\ SV3 + 0.4927\ STV1 + 3.1725$.

Cornell Index

Described in [Casale PN, et al; J. Am. Coll. Cardiol.; 1985; 6: 572-580] and [Casale PN, et al; Circulation; 1987; 75: 565-572]. The units are μV.

The Cornell index is the sum of the maximum R (or R') amplitude in aVL and the maximum of absolute S or absolute Q amplitude in V3. For females this value is increased by 600 μV.

Males: RaVL + |SV3|

Females: RaVL + |SV3| + 600

Cornell Product

Described in [Molloy TJ, et al; J. Am. Coll. Cardiol.; 1992; 20: 1180-1186-580] and [Okin PM, et al; J. Am. Coll. Cardiol.; 1995; 25: 417-423]. The units are μV.S.

This Cornell Product is defined as Cornell Index x QRS duration.

Sokolow-Lyon Index

Described in [Sokolow M, Lyon TP et al; Am. Heart J. 1949; 37: 161-186]. The units are μV.

The Sokolow-Lyon index is the sum of the maximum of absolute S or absolute Q amplitude in V1 and the maximum R (or R') amplitude in V5 or V6: |SV1| + maxRV5/6.

QRS Voltage Sum

Described in [Okin PM, et al; Hypertension 1998; 31: 937-942]. The units are μV.

QRS Voltage Sum is the sum of absolute values of QRS amplitudes in each lead.

QRS Voltage Product

Described in [Okin PM, et al; Hypertension 1998; 31: 937-942]. The units are μV.S.

The QRS Voltage Product is defined is the sum of the products of QRS duration and the absolute values of QRS amplitudes in each lead.

Label	Type	Commentary
p_duration	NUMBER	P duration.
lvmi_rautaharju	NUMBER	LVMI Rautaharju.
lvmi_f_huwez	NUMBER	LVMI F. Huwez.

Label	Type	Commentary
cornell_index	NUMBER	Cornell Index (μV).
cornell_product	NUMBER	Cornell Product (μV.S).
sokolow_lyon	NUMBER	Sokolow-Lyon Index (μV).
qrs_voltage_sum	NUMBER	QRS Voltage Sum (μV).
qrs_voltage_prod	NUMBER	QRS Voltage product (μV.S).

Glasgow Group Codes

The following table lists the Glasgow Group Codes output for each ECG.

Label	Type	Commentary
ecg_id	STRING (25)	The unique ID generated by the ECG management system.
statement_number	NUMBER(2)	Statement number within group.
text	STRING	Textual interpretation.

Minnesota Codes

The following table lists the Minnesota codes output for each ECG.

Label	Type	Commentary
ecg_id	STRING (25)	The unique ID generated by the ECG management system.
group_ident	STRING(2)	Group name / number.
statement_number	NUMBER(2)	Statement number within group.
non_site_specific	NUMBER(3)	Minnesota code for non-site specific groups.
anterolateral	NUMBER(3)	Minnesota code for anterolateral sites.
posterior	NUMBER(3)	Minnesota code for posterior sites.
anterior	NUMBER(3)	Minnesota code for anterior sites.