



ELSEVIER

Contents lists available at ScienceDirect

Data in brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Sport Database: Cardiorespiratory data acquired through wearable sensors while practicing sports



Agnese Sbröllini ^a, Micaela Morettini ^a, Elvira Maranesi ^b,
Ilaria Marcantoni ^a, Amnah Nasim ^a, Roberta Bevilacqua ^c,
Giovanni R. Riccardi ^b, Laura Burattini ^{a,*}

^a Cardiovascular Bioengineering Lab, Department of Information Engineering, Università Politecnica delle Marche, Ancona, Italy

^b Clinical Unit of Physical Rehabilitation, IRCCS INRCA, Ancona, Italy

^c Scientific Direction, IRCCS INRCA, Ancona, Italy

ARTICLE INFO

Article history:

Received 28 August 2019

Received in revised form 17 October 2019

Accepted 4 November 2019

Available online 12 November 2019

Keywords:

Sport acquisition

Exercise

Electrocardiogram

Heart-rate series

Breathing-rate series

ABSTRACT

Sport Database is a collection of 126 cardiorespiratory data, acquired through wearable sensors from 81 subjects while practicing 10 different sports. Each cardiorespiratory dataset consists of demographic info (gender, age, weight, height, smoking habit, alcohol consumption and weekly training rate), cardiorespiratory signals (electrocardiogram, heart-rate series, RR-interval series and breathing-rate series) and training notes. Demographic info was collected by survey. Cardiorespiratory signals were acquired through the chest strap BioHarness 3.0 by Zephyr. Eventually, training notes including the sport-dependent training protocol, were manually annotated. Sport Database may be useful to support: 1) the investigation of cardiorespiratory system adaptations to different types of physical exercise; 2) the development of automatic algorithms finalized to real-time health monitoring of athletes and preventive identification of subjects at increased risk of sport-related sudden cardiac death; and, 3) clinical testing of the BioHarness 3.0 by Zephyr. Further acquisitions could involve other

* Corresponding author. Department of Information Engineering, Università Politecnica delle Marche, via Brecce Bianche 12, 60131 Ancona, Italy.

E-mail address: l.burattini@univpm.it (L. Burattini).

sports, other cardiovascular signals and/or parameters, data from different biological systems, and other acquisition devices.

© 2019 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Specifications Table

Subject	Biomedical Engineering
Specific subject area	Cardiorespiratory data during sports
Type of data	Matlab Structures Text files
How data were acquired	BioHarness 3.0 by Zephyr (wearable sensor) and surveys
Data format	Raw and analyzed
Parameters for data collection	A total of 126 sets of cardiorespiratory data acquired from 81 athletes while practicing sports and consisting of demographic info (gender, age, weight, height, smoking habit, alcohol consumption and weekly training rate), cardiorespiratory signals (electrocardiograms, heart-rate series, electrocardiographic RR-interval series, breathing-rate series) and training notes.
Description of data collection	Demographic info was collected by survey and cardiorespiratory signals were recorded through the chest strap BioHarness 3.0 by Zephyr from athletes practicing 10 different sports (aerial silks, basketball, CrossFit, fitness, jogging, middle-distance running, running, soccer, tennis and Zumba). Acquisition protocol depended on practiced sport.
Data source location	Gyms or playing fields where the considered sports were performed and the Cardiovascular Bioengineering Lab (data owner and data storage location) of the Università Politecnica delle Marche, Ancona, Italy.
Data accessibility	With the article
Related research article	A. Agostinelli, M. Morettini, A. Sbröllini, E. Maranesi, L. Migliorelli, F. Di Nardo, S. Fioretti, L. Burattini, CaRiSMA 1.0: Cardiac Risk Self-Monitoring Assessment, Open Sports Sci. J. (2017). https://doi.org/10.2174/1875399X01710010179 [1].

Value of the Data

- Sport Database may be useful to investigate physiological and pathological adaptations of the cardiorespiratory system to different types of physical exercise.
- Sport Database may support the development of automatic algorithms finalized to real-time health monitoring of athletes and preventive identification of subjects at increased risk of sport-related sudden cardiac death.
- Sport Database may support clinical testing of the wearable sensor BioHarness 3.0 by Zephyr.
- Besides clinicians and biomedical engineers doing research on sport effects on athletes' health, personal trainers can benefit from these data to optimize training sessions from both health and performance points of view.
- Further acquisitions could involve other sports, other cardiovascular signals and/or parameters, data from different biological systems (for example the metabolic system and the motor system), and other acquisition devices.
- Additional value of these data consists in their usefulness to evaluate filtering procedures for cardiorespiratory signals, since acquisitions during exercise are affected by high levels of noise.

1. Data

Sport Database includes 126 cardiorespiratory datasets (CRD) from 81 subjects while performing 10 different sports: aerial silks, basketball, CrossFit, fitness, jogging, middle-distance running, running, soccer, tennis and Zumba (Table 1). Data are organized in a tree structure (Fig. 1). The main directory (SportDB) includes a folder for each sport (AER, BAS, CRO, FIT, JOG, MID, RUN, SOC, TEN and ZUM, respectively). Each sport folder contains a subfolder for each subject performing that sport (Sn, with $n = 1, 2 \dots$). Eventually, each subject subfolder contains a sub-subfolder for each acquisition performed by that subject (CRDm, with $m = 1, 2 \dots$). Each CRDm includes a demographic data file (Dem.txt), a

Table 1

Demographic data of Sport Database. Amount of missing data is reported in parenthesis. Overall values are computed excluding the missing data.

	Number of Subjects	Number of CRD	Gender M/F	Age (years)	Weight (kg)	Height (cm)	Smoking NO/YES	Alcohol consumption NO/SOMETIMES	Weekly training rate
AER	3	3	0/3 (0)	25 ± 3 (0)	53 ± 4 (0)	159 ± 1 (0)	-/- (3)	-/- (3)	-±- (3)
BAS	9	9	9/0 (0)	22 ± 4 (0)	74 ± 9 (0)	180 ± 5 (0)	1/8 (0)	0/9 (0)	4 ± 0 (0)
CRO	19	28	13/6 (0)	31 ± 7 (0)	71 ± 12 (0)	176 ± 7 (0)	9/10 (0)	4/15 (0)	4 ± 1 (0)
FIT	8	8	5/3 (0)	25 ± 5 (0)	71 ± 14 (0)	173 ± 7 (0)	4/4 (0)	4/4 (0)	4 ± 1 (0)
JOG	5	19	3/2 (0)	30 ± 14 (0)	63 ± 14 (3)	173 ± 8 (3)	3/- (2)	-/1 (4)	-±- (5)
MID	10	10	10/0 (0)	37 ± 16 (0)	70 ± 8 (0)	177 ± 3 (0)	9/1 (0)	2/8 (0)	4 ± 1 (0)
RUN	10	10	9/1 (0)	22 ± 3 (0)	70 ± 6 (0)	179 ± 7 (0)	5/5 (0)	1/9 (0)	3 ± 1 (0)
SOC	2	14	2/0 (0)	24 ± 1 (0)	67 ± 2 (0)	176 ± 1 (0)	-/- (2)	-/- (2)	-±- (2)
TEN	9	19	1/8 (0)	27 ± 11 (0)	60 ± 7 (1)	170 ± 6 (1)	8/1 (0)	0/9 (0)	3 ± 1 (1)
ZUM	6	6	1/5 (0)	35 ± 9 (1)	66 ± 17 (4)	174 ± 14 (4)	-/- (6)	-/- (6)	-±- (6)
Overall	81	126	53/28 (0)	30 ± 13 (1)	71 ± 21 (8)	170 ± 30 (8)	39/29 (13)	11/55 (15)	4 ± 1 (17)

AER = aerial silks; BAS = basketball; CRO = CrossFit; FIT = fitness; JOG = jogging; MID = middle-distance running; RUN = running; SOC = soccer; TEN = tennis; ZUM = Zumba.

cardiorespiratory data MATLAB structure (Data.mat) and a training note file (TrNote.txt). The demographic data file includes information about gender (male: 0; female: 1), age (years), weight (kg), height (cm), smoking habit (no: 0; yes: 1), alcohol consumption (no: 0; sometimes: 1) and weekly training rate (integer from 1 to 7); missing data are indicated with 'NA'. The cardiorespiratory data structure contains the recorded cardiorespiratory signals during the acquisition and includes four fields: Data.ECG, containing the raw electrocardiogram (ECG); Data.HR, containing the raw heart-rate (HR) series; Data.RR, containing the RR-interval series; and Data.BR containing the raw breathing-rate (BR) series. Characteristics of the cardiorespiratory signals (sampling frequency, amplitude range and data-loss index) are reported in Table 2. The training-notes file contains information about duration of the training phases during the acquisition and details about the sport-related acquisition protocol; acquisition phases annotated as 'none' indicate training phases not practiced by the subject.

2. Experimental design, materials, and methods

2.1. Data collection and acquisition

All subjects were supposed healthy (*i.e.* no previous history of cardiorespiratory diseases and not taking any drug) at the acquisition time. However, automatic analysis of acquired data by CaRiSMA software [1] suggested clinical evaluation to two subjects who were then diagnosed as affected by asymptomatic short QT syndrome (subject 5 practicing jogging) and paroxysmal atrial fibrillation (subject 9 practicing tennis). Demographic data of subjects are summarized in Table 1.

All subjects gave their informed consent prior to data collection and acquisitions, which were undertaken in compliance with the ethical principles of Helsinki Declaration and approved by the institutional expert committee.

Demographic data, cardiorespiratory signals and training notes of each CRD were collected during the same acquisition. Demographic data were collected by survey. Cardiorespiratory signals were

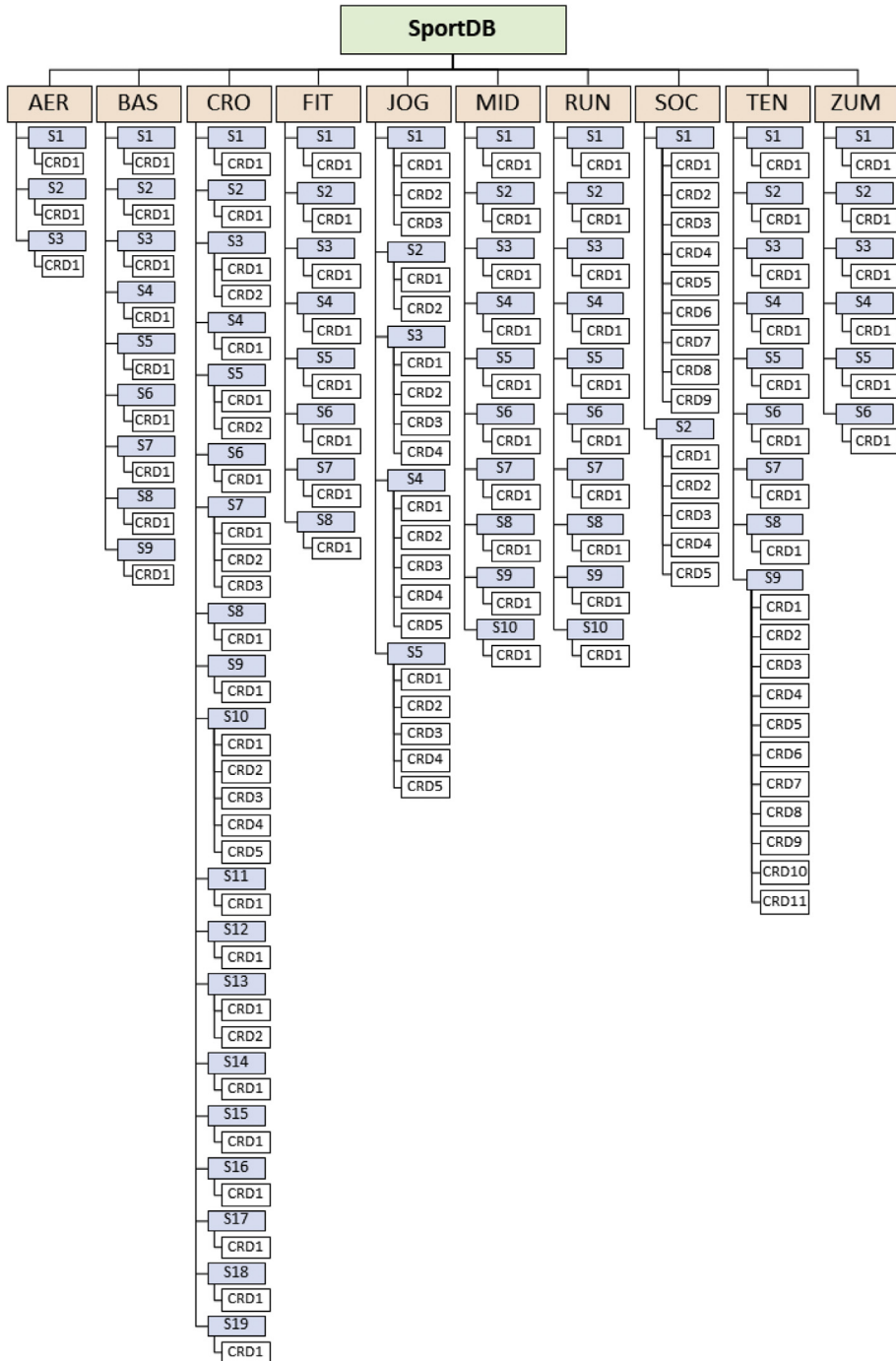


Fig. 1. Sport Database tree structure.

Table 2
Characteristics of the cardiorespiratory signals.

Signal	Sampling Frequency	Amplitude Range	Data Loss
ECG	250 Hz	0.25–15 mV	0 mV
HR	1 Hz	25–240 bpm	0 bpm
RR	1 Hz	250–2400 ms	Inf
BR	1 Hz	3–70 cpm	6553.5 cpm

ECG = electrocardiogram; HR = heart-rate series; RR = RR-interval series; BR = breathing-rate series.

recorded through the chest strap BioHarness 3.0 by Zephyr (www.zephyranywhere.com), a reliable wearable device [1–6] that directly records the ECG (mV; raw data) and automatically computes the HR series (bpm; processed data) and the BR series (cpm; processed data).

The characteristics of these signals are reported in Table 2. Before each acquisition, the strap was slightly moistened in order to optimize electrical conductivity. Then, the device sensor was positioned under the left arm, as suggested by guidelines. The RR-interval series (ms) were indirectly computed from HR series:

$$RR = \frac{60}{HR} \cdot 1000 \quad (1)$$

During each acquisition, all information about duration of the training phases and details about the sport-related acquisition protocol were manually annotated in the training note file.

2.2. Acquisition protocols

A specific acquisition protocol was defined for each sport. Each acquisition protocol includes several phases, the starting and the duration of which was varying and measured using a stopwatch and reported in the training note file.

2.2.1. Aerial Silks Protocol

The aerial silks protocol includes three phases: an initial phase of resting, a phase of exercise and a final phase of recovery. During the resting phase the subject sits courtside. During the exercise phase he/she performs aerial silks exercises. Finally, during the recovery phase he/she sits courtside again.

2.2.2. Basketball Protocol

The basketball protocol includes three main phases: an initial phase, a central phase and a final phase. The initial phase is a warm-up phase, composed by four combinations of exercises: layups; one-hand passes and shots; dribbles, passes and shots; and dribbles, passes, shots and defense. The central phase is a simulation phase, composed by five offense-defense exercises: a 2-men offense vs a 1-man defense; a 3-men offense vs a 2-men defense; a 3-men offense vs a 3-men defense; a 4-men offense vs a 4-men defense; and a free 4-men vs 4-men match. The final phase is a 5-men vs 5-men match. All three main phases could be interrupted by short resting phases during which the subject sits courtside.

2.2.3. CrossFit Protocol

The CrossFit protocol includes three phases: a warm-up phase, a skill phase and a workout-of-the-day phase. The third phase is mandatory, while the other two may not be performed. Each phase may involve different types of exercises varying from subject to subject.

2.2.4. Fitness Protocol

The fitness protocol includes two main phases: an exercise phase and a recovery phase. The exercise phase includes a series of exercises varying from subject to subject and could be interrupted by short resting phases. After training, there is the recovery phase during which the subject sits courtside.

2.2.5. Jogging Protocol

The jogging protocol includes three main phases: an initial phase of resting, a phase of exercise and a final phase of recovery. During the resting phase the subject sits courtside. During the exercise phase the subject performs a freely chosen jogging training that can be done on a treadmill or outdoor. Finally, during the recovery phase the subject sits courtside again.

2.2.6. Middle-Distance Running Protocol

The middle-distance running protocol includes three phases: an initial phase of resting, a phase of exercise and a final phase of recovery. During the resting phase the subject sits courtside. During the exercise phase the subject runs for 2 km in a standard 400 m track and independently increases his/her speed every 200 m (50% of the track), thus performing a standard Conconi's test [7]. Finally, during the recovery phase the subject sits courtside again.

2.2.7. Running Protocol

The running protocol is called Around Ancona [8] and it includes a close 6.1 Km route around the city of Ancona (Fig. 2). The starting and ending point of the route is located at the Monumento dei Caduti memorial. The protocol includes four phases with different slopes: an initial flat phase, an uphill phase, a downhill phase and a final flat phase. The initial flat phase (Fig. 2-blue line) is 1.3 km long with a 0% slope; the uphill phase (Fig. 2-red line) is 1.2 km long with a +6.8% slope; the downhill phase (Fig. 2-purple line) is 1 km long with a -7.2% slope; the final flat phase (Fig. 2-green line) is 2.6 km long with a 0% slope.

2.2.8. Soccer Protocol

The soccer protocol includes three phases: an initial phase of resting, a phase of exercise and a final phase of recovery. During the resting phase the subject sits courtside. During the exercise phase the subject plays soccer. Finally, during the recovery phase the subject sits courtside again.

2.2.9. Tennis Protocol

The tennis protocol includes three phases: an initial phase of resting, a phase of exercise and a final phase of recovery. During the resting phase the subject sits courtside. During the exercise phase the subject plays tennis. Finally, during the recovery phase the subject sits courtside again.

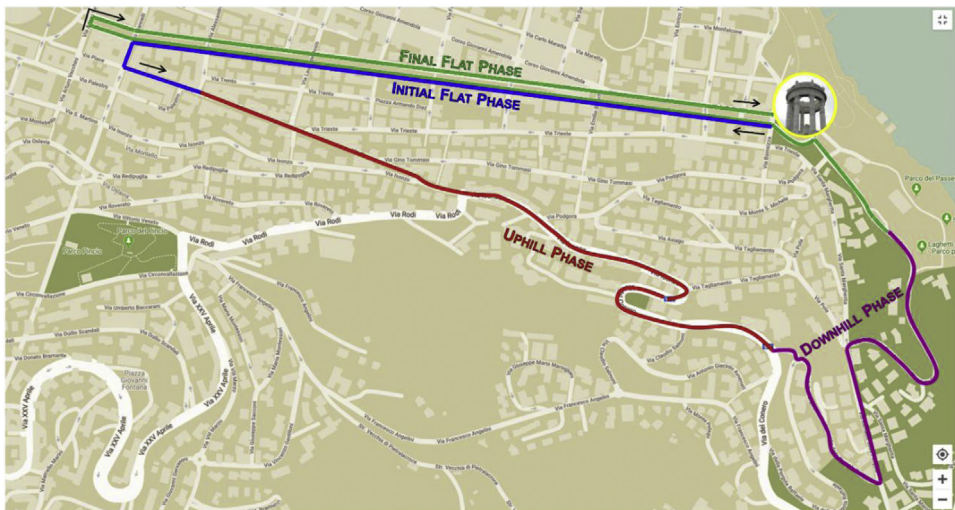


Fig. 2. Around Ancona route. The route starts and ends at the "Monumento dei Caduti" and it is composed of four phases: an initial flat phase (blue line), an uphill phase (red line), a downhill phase (purple line) and a final flat phase (green line).

2.2.10. Zumba Protocol

The Zumba protocol includes three phases: an initial phase of resting, a phase of exercise and a final phase of recovery. During the resting, the subject sits courtside. During the exercise phase the subject performs Zumba. Finally, during the recovery phase the subject sits courtside again.

Acknowledgments

Authors would like to acknowledge Consortium GARR (<https://www.garr.it>) for financing the project entitled “Sistema di Monitoraggio Cardiovascolare in-Cloud per la Prevenzione della Morte Cardiaca Improvvisa da Sport” (System for in-Cloud Cardiovascular Monitoring for Prevention of Sport-Related Sudden Cardiac Death).

Conflict of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.dib.2019.104793>.

References

- [1] A. Agostinelli, M. Morettini, A. Sbröllini, E. Maranesi, L. Migliorelli, F. Di Nardo, S. Fioretti, L. Burattini, CaRiSMA 1.0: cardiac risk self-monitoring assessment, *Open Sports Sci. J.* (2017), <https://doi.org/10.2174/1875399X01710010179>.
- [2] G. Nazari, J.C. MacDermid, K.E. Sinden, J. Richardson, A. Tang, Reliability of zephyr bioharness and fitbit charge measures of heart rate and activity at rest, during the modified Canadian aerobic fitness test, and recovery, *J. Strength Cond. Res.* (2019), <https://doi.org/10.1519/JSC.0000000000001842>.
- [3] M. Al Ahmad, S. Ahmed, Piezologist: a novel wearable piezoelectric-based cardiorespiratory monitoring system, in: 2018 IEEE Int. Conf. Innov. Intell. Syst. Appl. INISTA, 2018, <https://doi.org/10.1109/INISTA.2018.8466275>.
- [4] J. de Bruijn, H. van der Worp, M. Korte, A. de Vries, R. Nijland, M. Brink, Sport-specific outdoor rehabilitation in a group setting: do the intentions match actual training load? *J. Sport Rehabil.* (2018) <https://doi.org/10.1123/jsr.2016-0009>.
- [5] D. Nepi, A. Sbröllini, A. Agostinelli, E. Maranesi, M. Morettini, F. Di Nardo, S. Fioretti, P. Pierleoni, L. Pernini, S. Valenti, L. Burattini, Validation of the heart:rate signal provided by the zephyr BioHarness 3.0, in: 2016 Comput. Cardiol. Conf., 2016, <https://doi.org/10.22489/CinC.2016.106-358>.
- [6] C. Massaroni, A. Nicolò, D. Lo Presti, M. Sacchetti, S. Silvestri, E. Schena, Contact-based Methods for Measuring Respiratory Rate, *Sensors (Switzerland)*, *Sensors (Switzerland)* (2019), <https://doi.org/10.3390/s19040908>.
- [7] F. Conconi, M. Ferrari, P.G. Ziglio, P. Droghetti, L. Codeca, Determination of the anaerobic threshold by a noninvasive field test in runners, *J. Appl. Physiol.* (2017), <https://doi.org/10.1152/jappl.1982.52.4.869>.
- [8] A. Sbröllini, G. Caraceni, A. Nasim, I. Marcantoni, M. Morettini, A. Belli, P. Pierleoni, L. Burattini, Self-monitoring of cardiac risk while running around Ancona, in: ISCT, 2019.