# Sport DB 2.0: a New Database of Data Acquired by Wearable and Portable Devices while Practicing Sport

Sofia Romagnoli[1], Agnese Sbrollini[1], Antonio Nocera[1], Micaela Morettini[1], Ennio Gambi[1], Danilo Bondi[2], Tiziana Pietrangelo[2], Vittore Verratti[2], Laura Burattini[1,*]

[1]Università Politecnica delle Marche, Ancona, Italy
[2]Università degli studi 'G. d'Annunzio' Chieti-Pescara, Chieti, Italy

**Abstract**

*Sport DB 2.0 is a collection of 156 cardiorespiratory datasets, acquired through wearable sensors and portable devices from 130 subjects while practicing 11 different sports during training and competition. Each dataset consists of demographic data (sex, age, weight, height, smoking habit, alcohol consumption, caffeine consumption, weekly training rate, presence of diseases and dietary supplement consumption), cardiorespiratory signals (electrocardiogram, heart-rate series, RR-interval series, and/or breathing-rate series), and training note data (sport-dependent training protocol). Cardiorespiratory signals were acquired through the BioHarness 3.0 by Zephyr, the KardiaMobile by AliveCor, the Kardia 6L by AliveCor, the Polar M400 by Polar, and heart-rate sensor H7 by Polar, on the playing field or gym following a specific acquisition protocol for each sport. Sport DB 2.0 may be useful to support research activity finalized to investigate the cardiorespiratory pathophysiological mechanisms triggered by sport, to develop automatic algorithms for monitoring athletes' health while practicing sports, to validate the reliability of wearable sensors and portable devices in sport, and to develop data analytics techniques and artificial intelligence applications to support sport sciences.*

## 1. Introduction

Sport is recommended for its beneficial effects on health but it rarely may trigger acute cardiac events [1-2]. Understanding the hidden mechanisms occurring during sport is crucial [1-2]; thus, cardiorespiratory data in athletes are essential to support such research activity.

Nowadays, athletes are used to monitor their exercise dose and intensity through wearable sensors and portable devices. These technologies were proven to be reliable in the recording of cardiorespiratory signals [2], which are the electrocardiogram (ECG), the heart rate (HR) series, and the breathing rate (BR) series. Consequently, research activity on data acquired by wearable and portable devices may provide some insights into athletes' health.

In 2019, Sbrollini et al. [1] published Sport Database, a collection of cardiorespiratory data acquired through the chest band BioHarness 3.0 by Zephyr while practicing sports. Sport Database includes 126 cardiorespiratory datasets from 81 subjects while performing 10 different sports [1] and it was used to investigate normative values of athlete's electrocardiogram [2], to develop automatic algorithms for monitoring athletes' health [3-5], to validate the reliability of wearable and portable devices in sport [6-8], and to assess sport analytics applications [9]. Still, a recent scoping review on wearable sensors and portable devices highlighted the need of other open-source databases of cardiorespiratory signals acquired while practicing sports [10]. Thus, the present work aims to present Sport DB 2.0, a database including a new set of cardiorespiratory data acquired by wearable sensors and portable devices during 11 different sports.

## 2. Database organization

Sport DB 2.0 includes 156 cardiorespiratory datasets (CRD) from 130 subjects while performing 11 different sports, either during the whole session or before-after (Figure 1). Data are organized in the main 'SportDB2' directory, including 11 folders (one for each sport), which are American football (AMF), athletics (ATH), basketball (BAS), cardio-fitness (CFIT), cycling (CYC), futsal (FUT), speed skating (SKA), running (RUN), soccer (SOC), trekking (TRE), and volleyball (VOL). The BAS and SOC folders are organized into two subfolders (BAS1, BAS2) and three subfolders (SOC1, SOC2, SOC3), respectively. Each sport folder contains a subfolder for each subject performing that specific sport and each subject subfolder contains a subfolder for each CRD.

Each CRD folder includes a demographic data file ('Dem.txt'), a MATLAB data file ('Data.mat'), and a training note file ('TrNote.txt'). The demographic data file includes information about sex (male: 0; female: 1), age (years), weight (kg), height (cm), smoking habit (no: 0;

yes: 1), alcohol consumption (no: 0; yes: 1; if 24 before the training), caffeine consumption (if 6h before the training), weekly training rate (times per week), presence of diseases, and dietary supplement consumption. Missing data are indicated with 'NA'. The MATLAB data file includes cardiorespiratory data acquired through wearable sensors or portable devices, and its format depends on acquisition system. In case of data acquired by BioHarness 3.0, the MATLAB data file includes 'Data.ECG', containing raw ECG (sample rate, sf = 250 Hz), 'Data.HR', containing raw HR series (sf = 1 Hz), 'Data.RR', containing the derived RR-interval series (sf = 1 Hz) [1], and 'Data.BR' containing the raw BR series (sf = 1 Hz). In case of data acquired by KardiaMobile or Kardia 6L, the MATLAB data file includes 'Data.ECG' containing raw ECG (sf = 300 Hz). In case of data acquired by Polar M400 or heart sensor H7, the MATLAB data file includes 'Data.HR', containing the raw HR series (sf = 1 Hz). Training note file contains details of acquisition protocols.

## 3. Acquisition protocols

All subjects were healthy (no previous history of diseases and not taking any drug) at the acquisition time, except for one subject in SOC3, who was affected by atrial stenosis and tachycardia of unknown origin. All subjects gave their informed consent prior to data acquisition, which was undertaken in compliance with the ethical principles of the Helsinki Declaration and approved by the institutional expert committee.

Demographic data were collected by survey, and they are summarized in Table 1.

Cardiorespiratory signals were recorded through wearable sensors or portable devices, that are the BioHarness 3.0 by Zephyr, KardiaMobile by AliveCor, Kardia 6L by AliveCor, Polar M400 by Polar, and heart-rate sensor H7 by Polar.

A specific acquisition protocol was defined for each sport (details are summarized in Table 2) and reported in the training note file. All acquisition protocols include three main phases: a resting phase (at least 5 minutes at the courtside), an exercise phase (free duration), and a recovery phase (at least 5 minutes; it coincides with stretching). Specifically, each acquisition protocol may include several different sport-dependent phases, which starting and duration were measured using a stopwatch.

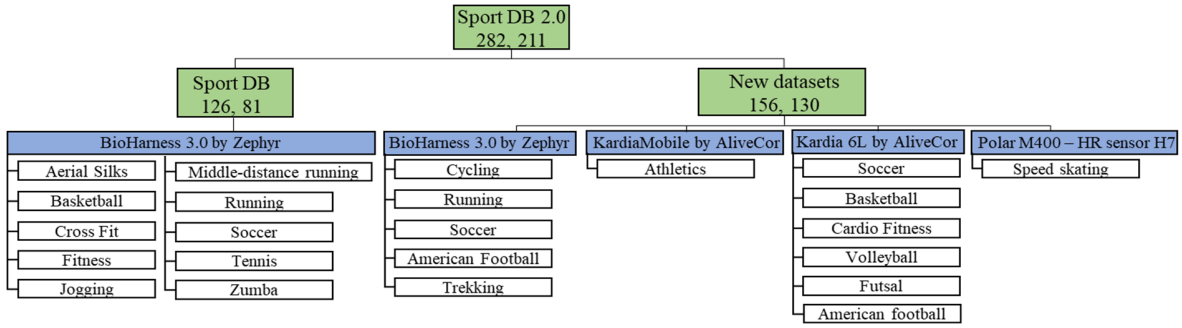


Figure 1. Structure of Sport DB 2.0 database. N,S indicates the number of datasets and the number of subjects.

Table 1. Summary of demographic data acquired in Sport DB 2.0.

| Sport | N. of subjects | N. of CRD | Sex M/F | Age (years) | Weight (Kg) | Height (cm) | Smoking no/yes | Alcohol no/yes | Weekly training rate |
|---|---|---|---|---|---|---|---|---|---|
| AMF | 15 | 15 | 15/0 | 24±4 | 95±15 | 182±7 | 9/6 | 0/15 | 3±0 |
| ATH | 10 | 30 | 1/9 | 18±2 | 59±5 | 169±9 | 10/0 | 5/5 | 5±1 |
| BAS1 | 11 | 11 | 11/0 | 19±1 | 78±6 | 187±5 | 7/3 | 1/10 | 5±1 |
| BAS2 | 22 | 22 | 11/11 | 24±4 | 81±19 | 185±12 | 15/7 | 4/18 | 5±2 |
| CFIT | 6 | 6 | 3/3 | 32±13 | 71±20 | 172±13 | 4/2 | 1/5 | 4±2 |
| CYC | 12 | 12 | 10/2 | 33±16 | –±– | –±– | 12/0 | –/– | –±– |
| FUT | 7 | 7 | 7/0 | 22±3 | 69±8 | 177±5 | 7/3 | 1/6 | 4±0 |
| SKA | 1 | 4 | 0/1 | 22 | 58 | 170 | 1/0 | 0/1 | 5 |
| RUN | 10 | 10 | 9/1 | 21±1 | 69±6 | 178±6 | 6/4 | 1/9 | 4±1 |
| SOC1 | 10 | 10 | 10/0 | 29±7 | 76±9 | 180±5 | 7/3 | 0/10 | 3±1 |
| SOC2 | 9 | 14 | 9/0 | 31±8 | 74±9 | 177±5 | 7/3 | 0/9 | 3±0 |
| SOC3 | 1 | 1 | 1/0 | 23 | 77 | 178 | 1/0 | 0/1 | 3 |
| TRE | 3 | 11 | 0/3 | 25;28;26 | 52;63;59 | 161;163;161 | 3/0 | –/– | –±– |
| VOL | 13 | 13 | 0/13 | 23±3 | 71±8 | 173±5 | 11/2 | 0/13 | 4±1 |

Table 2. Acquisition protocols.

| Sport | Device | Exercise phase content | Time of acquisition* |
|---|---|---|---|
| AMF | BioHarness 3.0 by Zephyr<br>Kardia 6L by AliveCor | Mobility, linear speed, technical exercises, and game simulation. | Acquisition by the BioHarness 3.0 lasts for the entire duration of the training.<br>Four 30-second acquisitions by Kardia 6L were performed at (1) resting, (2) 3-min after the acquisition (1); (3) 0-min, (4) 5-min after the end of training |
| ATH | KardiaMobile by AliveCor | *Strength training*: warm-up, 6 series of mid squat jump and step jump;<br>*Speed training*: warm-up, 2 sprint races of 200m and one sprint race of 150m;<br>*Competition*: warm-up and 1 sprint race of 200m. | Six 30-second acquisitions were performed at (1) resting, (2) post warm-up; (3) 0-min, (4) 5-min, (5) 10-min, and (6) 15-min after the end of training.<br>Some additional acquisitions were performed during strength training and speed training in the middle of the exercise. |
| BAS - BAS1 | Kardia 6L by AliveCor | Warm-up, technical exercises, and match simulation. | Six 30-second acquisitions were performed at (1) resting, (2) post warm-up; (3) 0-min, (4) 5-min, (5) 10-min, and (6) 15-min after the end of training. |
| BAS - BAS2 | Kardia 6L by AliveCor | Warm-up and match. | Six 30-second acquisitions were performed at (1) resting, (2) post warm-up; (3) 0-min, (4) 5-min, (5) 10-min, and (6) 15-min after the end of training. |
| CFIT | Kardia 6L by AliveCor | *Cardio training*: running at sustained speed;<br>*Functional training:* functional exercises following TABATA system. | Six 30-second acquisitions were performed at (1) resting, (2) post warm-up; (3) 0-min, (4) 5-min, (5) 10-min, and (6) 15-min after the end of training.<br>Some additional acquisitions were performed during functional training at the end of each exercise. |
| CYC | BioHarness 3.0 by Zephyr | 13,9 km loop with 473 m of elevation. | Acquisition lasts for the entire duration of the training |
| FUT | Kardia 6L by AliveCor | Warm-up, technical exercises, and match simulation. | Six 30-second acquisitions were performed at (1) resting, (2) post warm-up; (3) 0-min, (4) 5-min, (5) 10-min, and (6) 15-min after the end of training. |
| SKA | Polar M400 by Polar;<br>HR sensor H7 by Polar | Warm-up, exercises with different duration and power. | Acquisition lasts for the entire duration of the training |
| RUN | BioHarness 3.0 by Zephyr | Around Ancona protocol [11]. | Acquisition lasts for the entire duration of the training |
| SOC-SOC1 | Kardia 6L by AliveCor | Warm-up, strength exercises, resistance exercises, speed exercises, and match simulation. | Six 30-second acquisitions were performed at (1) resting, (2) post warm-up; (3) 0-min, (4) 5-min, (5) 10-min, and (6) 15-min after the end of training. |
| SOC-SOC2 | Kardia 6L by AliveCor | Warm-up, strength exercises, resistance exercises, speed exercises, technical exercises, and match simulation. | Six 30-second acquisitions were performed at (1) resting, (2) post warm-up; (3) 0-min, (4) 5-min, (5) 10-min, and (6) 15-min after the end of training. |
| SOC3 | BioHarness 3.0 by Zephyr | Strength exercises, resistance exercises, and technical trainings. | Acquisition lasts for the entire duration of the training |
| TRE | BioHarness 3.0 by Zephyr | Trek up to 4,556m of altitude. | Acquisition lasts for the entire duration of the training |
| VOL | Kardia 6L by AliveCor | Warm-up, technical exercises, match simulation. | Six 30-second acquisitions were performed at (1) resting, (2) post warm-up; (3) 0-min, (4) 5-min, (5) 10-min, and (6) 15-min after the end of training. |

*: details regarding the time of acquisitions are reported in the training note file

## 4. Discussion

Sport DB 2.0 is a novel database, the construction of which was inspired by Sport Database [1]. Differently, Sport DB 2.0 includes additional sport activities. Moreover, data in Sport Database was uniquely acquired by BioHarness 3.0, while Sport DB 2.0 includes data acquired by different wearable sensors and portable devices, produced by different companies.

Sport DB 2.0 may be useful to support research activity on pathophysiological cardiorespiratory mechanisms triggered by sport activity. Clinically, it may support the development of new automatic algorithms for athletes' health monitoring. These tools may be used by coaches to optimize athletes' training/performance and by clinicians to monitor athletes' health. Technically, Sport DB 2.0 may support the quality assessment of data acquired by wearable sensors and portable devices in sport applications. Indeed, such data are usually characterized by high levels of interferences. The database may support the design of novel signal processing procedures for cardiorespiratory data denoising. Finally, Sport DB 2.0 may support the development of the novel techniques of data analytics and artificial intelligence applied to sport science and exercise and network physiology.

Utility of data collected on Sport DB 2.0 was already proven in the literature. CYC dataset was used to investigate normative values of athletes' ECG [2] and TRE dataset was used to assess cardiorespiratory mechanisms during high-altitude activity [13]. Finally, of note, ATH dataset is one of the rarely collection of data acquired during competition. Indeed, this rare dataset was used to investigate the autonomous nervous system role by means of heart-rate variability indices in athletes under high levels of physical and psychological stressors [12].

## 5. Conclusion

Sport DB 2.0 represents a valid database to support the bioengineering research in sport: from the study of cardiorespiratory mechanisms associated to sport activity to the development of innovative signal processing algorithms.

## Acknowledgments

The research presented in this manuscript is supported by the project "Chaalenge" (CUP B39J2200305000).

## References

[1] A. Sbrollini et al., "Sport Database: Cardiorespiratory data acquired through wearable sensors while practicing sports," *Data Brief.*, vol. 27, no. 104793, Nov., 2019.

[2] S. Romagnoli et al., "Initial investigation of athletes' electrocardiograms acquired by wearable sensors during the pre-exercise phase," *TOBEJ*, vol. 15, pp. 37–44, July, 2021.

[3] S. Romagnoli; A. Sbrollini; I. Marcantoni; M. Morettini; L. Burattini, "Sport?Sicuro! A Graphical User Interface for Continuous Cardiovascular Monitoring while Playing Sport Based on Heart Rate and Heart-Rate Variability," *2022 Computing in Cardiology (CinC)*, pp. 1-4, Sept., 2022.

[4] S. Romagnoli; I. Marcantoni; K. Campanella; A. Sbrollini; M. Morettini; L. Burattini, "Ensemble Empirical Mode Decomposition for Efficient R-Peak Detection in Electrocardiograms Acquired by Portable Sensors During Sport Activity," *IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, June, 2021.

[5] S. Romagnoli et al., "Signal Processing for Athletic Cardiovascular Monitoring with Wearable Sensors: Fully Automatic Detection of Training Phases from Heart Rate Data," *2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pp. 1491–1494, Dec., 2021.

[6] S. Andjelic, C. Doyle and G. Hossain, "Cybersecurity Risk with Wearable Technology in Sports: Why Should We Care?," *2022 IEEE 19th International Conference on Smart Communities: Improving Quality of Life Using ICT, IoT and AI (HONET)*, pp. 160-165, Dec., 2022.

[7] S. Miccinilli et al., "Use of wearable systems for the detection of chest-abdominal wall movement aimed at respiratory monitoring in sport: a scoping review on available data," *Journal of biological regulators and homeostatic agents,* vol. 34, pp. 87–96, Sep-Oct 2020.

[8] B. Širaiy, V. Ilić, L. Toskić, "Usability of Wireless ECG Body Sensor for Cardiac Function Monitoring During Field Testing," *Measurement Science Review*, vol. 21, pp. 55–60, May, 2021.

[9] A. Rajšp, I. Fister, "A Systematic Literature Review of Intelligent Data Analysis Methods for Smart Sport Training," *Applied Sciences*, vol. 10, no. 9, p. 3013, Apr., 2020.

[10] S. Romagnoli, F. Ripanti, M. Morettini, L. Burattini, A. Sbrollini, "Wearable and Portable Devices for Acquisition of Cardiac Signals while Practicing Sport: A Scoping Review," *Sensors*, vol. 23, no. 6, p. 3350, Mar., 2023.

[11] A. Sbrollini et al., "Self-Monitoring of Cardiac Risk while Running Around Ancona," *2019 IEEE 23rd International Symposium on Consumer Technologies (ISCT)*, pp. 1–4, June, 2019.

[12] S. Romagnoli, A. Sbrollini, M. Morettini, L. Burattini, "Symbolic Analysis of Heart-Rate Variability during Training and Competition in Short Distance Running," *2023 IEEE 36th International Symposium on Computer-Based Medical Systems (CBMS)*, pp. 585–588, June, 2023.

[13] A. Sbrollini et al., "Segmented-Beat Modulation Method-Based Procedure for Extraction of Electrocardiogram-Derived Respiration from Data Acquired by Wearable Sensors During High-Altitude Activity," *2022 Computing in Cardiology (CinC)*, pp. 1–4, Sept., 2022.

Address for correspondence:
Laura Burattini.
Department of Information Engineering, Università Politecnica delle Marche, Via brecce Bianche 12, 60131, Ancona, Italy.
l.burattini@univpm.it