

# Author Contributions Checklist Form

This form documents the artifacts associated with the article (i.e., the data and code supporting the computational findings) and describes how to reproduce the findings.

## Part 1: Data

☐ This paper **does not** involve analysis of external data (i.e., no data are used or the only data are generated by the authors via simulation in their code).

☒ I certify that the author(s) of the manuscript have legitimate access to and permission to use the data used in this manuscript.

## Abstract

The clinical dataset is composed of a total of 711 German, Austrian and Italian patients suffering from heart failure in different stages, in hospital facility due to either an acute hospitalization or an ambulatory visit, released and followed-up for a period of 6 months. Patients were enrolled in two distinct clinical trials. The related protocol and amendments were approved by the institutional review board at each participating center, and the trials were conducted in accordance with the principles of the Declaration of Helsinki, Good Clinical Practice guidelines, and local and national regulations. Written informed consent was provided by all patients before any study-related procedures were performed.

## Availability

☒ Data **are** publicly available

☐ Data **cannot be made** publicly available

If the data are publicly available, see the *Publicly available data* section. Otherwise, see the *Non-publicly available data* section, below.

## Publicly available data

☐ Data are available online at:

☒ Data are available as part of the paper's supplementary material.

☐ Data are publicly available by request, following the process described here:

☐ Data are or will be made available through some other mechanism, described here:

## Non-publicly available data

Discussion of lack of publicly available data:

## Description

### File format(s)

- ☐ CSV or other plain text:
- ☐ Software-specific binary format (.Rda, Python pickle, etc.):
- ☐ Standardized binary format (e.g., netCDF, HDF5, etc.):
- ☒ Other (described here):

Microsoft Excel (the extension of the file is .xlsx)

### Data dictionary

- ☒ Provided by the authors in the following file(s): Reproducibility.pdf (in Supplementary Materials)
- ☐ Data file(s) is (are) self-describing (e.g., netCDF files)
- ☐ Available at the following URL:

### Additional information (optional)

## Part 2: Code

### Abstract

The codes use the provided clinical dataset to generate enhanced datasets and the related statistics as described in the paper.

### Description

#### Code format(s)

- ☒ Script files
- ☐ R   ☐ Python   ☒ Matlab
  - ☐ Other:
- ☐ Package
- ☐ R   ☐ Python   ☐ MATLAB toolbox
  - ☐ Other:
- ☐ Reproducible report
- ☐ R Markdown   ☐ Jupyter notebook
  - ☐ Other:
- ☐ Shell script
- ☐ Other (described here):

### Supporting software requirements

Version of primary software used

MATLAB release R2020a

Libraries and dependencies used by the code

Statistics and Machine Learning Toolbox

### Supporting system/hardware requirements (optional)

### Parallelization used

- ☒ No parallel code used
- ☐ Multi-core parallelization on a single machine/node  
Number of cores used:
- ☐ Multi-machine/multi-node parallelization  
Number of nodes and cores used:

### License

- ☒ MIT License (default)
- ☐ BSD
- ☐ GPL v3.0
- ☐ Creative Commons
- ☐ Other (described here):

### Additional information (optional)

## Part 3: Reproducibility workflow

### Scope

The provided workflow reproduces:

- ☒ Any numbers provided in text in the paper
- ☐ All tables and figures in the paper
- ☒ Selected tables and figures in the paper, as explained and justified here:

The codes use the provided clinical dataset to generate enhanced datasets and the related statistics as described in the paper.

Note that the codes use the random number generator of MATLAB, so the results, the Figures 2B and 2C, and the statistics can be different at each script run.

In addition, the tables are obtained via the MATLAB® Classification Learner application (all default settings were unchanged) and via the MATLAB® Regression Learner application (all default settings were unchanged).

### Workflow details

#### Format(s)

- ☐ Single master code file
- ☐ Wrapper (shell) script(s)
- ☐ Self-contained R Markdown file, Jupyter notebook, or other literate programming approach
- ☐ Text file (e.g., a readme-style file) that documents workflow
- ☐ Makefile
- ☒ Other (more detail in 'Instructions' below)

#### Instructions

The details to run the scripts are in the file Reproducibility.pdf (in Supplementary Materials).

### Expected run-time

Approximate time needed to reproduce the analyses on a standard desktop machine:

- ☐ <1 minute
- ☐ 1-10 minutes
- ☒ 10-60 minutes

☐ 1-8 hours

☐ >8 hours

☐ Not feasible to run on a desktop machine, as described here:

Additional documentation (optional)

Notes (optional)