Package 'ABPS'

October 18, 2018

Title The Abnormal Blood Profile Score to Detect Blood Doping

Version 0.3

Date 2018-10-18

Description An implementation of the Abnormal Blood Profile Score (ABPS,
part of the Athlete Biological Passport program of the World Anti-Doping
Agency), which combines several blood parameters into a single score in
order to detect blood doping (Sottas et al. (2006)
<doi:10.2202/1557-4679.1011>). The package also contains functions to
calculate other scores used in anti-doping programs, such as the
OFF-score (Gore et al. (2003) <http://www.haematologica.org/content/88/3/333>),
as well as example data.

Imports kernlab

License GPL (>= 2)

Encoding UTF-8

LazyData true

RoxygenNote 6.0.1

Suggests testthat

NeedsCompilation no

Author Frédéric Schütz [aut, cre],
Alix Zollinger [aut]

Maintainer Frédéric Schütz <schutz@mathgen.ch>

Repository CRAN

Date/Publication 2018-10-18 12:30:10 UTC

R topics documented:

ABPS   *A function for calculating the Abnormal Blood Profile Score*

Description

The ABPS function computes the Abnormal Blood Profile Score from seven haematological markers. Higher values of this composite score are associated with a higher likelihood of blood doping.

Usage

```
ABPS(haemdata = NULL, HCT = NULL, HGB = NULL, MCH = NULL, MCHC = NULL,
  MCV = NULL, RBC = NULL, RETP = NULL)
```

Arguments

haemdata	a vector or data frame containing (at least) the 7 haematological variables, either with the same names as the parameters below, or (not recommended) without names but in the same order as the parameters.
HCT	haematocrit level [%]
HGB	haemoglobin level [g/dL]
MCH	mean corpuscular haemoglobin [pg]
MCHC	mean corpuscular haemoglobin concentration [g/dL]
MCV	mean corpuscular volume [fL]
RBC	red blood cell count [10^6/uL]
RETP	reticulocytes percent [%]

Details

The ABPS uses the seven haematological variables (HCT, HGB, MCH, MCHC, MCV, RBC, RETP) in order to obtain a combined score. This score is more sensitive to doping than the individual markers, and allows the detection of several types of blood doping using a single score.

The combined score is based on two classification techniques, a naive Bayesian classifier and an SVM (Support Vector Machine). The two models were trained using a database of 591 blood profiles (including 402 control samples from clean athletes and 189 samples of athletes who abused of an illegal substance); the two scores were then combined using ensemble averaging to obtain the final ABPS score.

The ABPS is part of the Athlete Biological Passport program managed by the World Anti-Doping Agency. While it is not a primary marker of doping, it has been used as corroborative evidence (see e.g. https://jurisprudence.tas-cas.org/Shared%20Documents/2773.pdf)

Value

a vector containing the ABPS score(s). Scores between 0 and 1 indicate a possible suspicion of doping; a score above 1 should only be found in 1 in 1000 male athletes.

Note

The values for the markers can be specified using either a data frame containing (at least) the 7 haematological variables, or using seven named parameters, but not both at the same time.

The calculation of the ABPS depends on two sets of parameters, for the two machine learning techniques (naive Bayesian classifier and Support Vector Machine), which are provided in the package.

Each parameter must be in a prespecified range; parameters outside this range are constrained to the min (respectively max) values.

Note that several versions of the ABPS were developed (including several different combinations of parameters). The version provided in this package provides the same results as the WADA version included in their ADAMS database. However, some values calculated with other versions of the software have also been distributed (see the help page for the `blooddoping` dataset for an example).

References

Sottas, P.E., N. Robinson, S. Giraud, et al., Statistical classification of abnormal blood profiles in athletes. Int J Biostat, 2006. 2(1): p. 1557-4679.

https://jurisprudence.tas-cas.org/Shared%20Documents/2773.pdf

Examples

```
ABPS(HCT=43.2, HGB=14.6, MCH=31.1, MCHC=33.8, MCV=92.1, RBC=4.69, RETP=0.48)
ABPS(data.frame(HCT=43.2, HGB=14.6, MCH=31.1, MCHC=33.8, MCV=92.1, RBC=4.69, RETP=0.48))
ABPS(c(43.2, 14.6, 31.1, 33.8, 92.1, 4.69, 0.48))
data(blooddoping); ABPS(blooddoping)
data(bloodcontrol); ABPS(bloodcontrol)
```

---

bloodcontrol *Blood samples from different individuals.*

Description

A dataset containing the result of the analysis of 13 blood samples from different individuals.

Usage

```
bloodcontrol
```

Format

A data frame with 13 rows and 12 haematological variables:

HCT haematocrit [%]

HGB haemoglobin [g/dL]

IRF immature reticulocyte fraction [%]

MCH mean corpuscular haemoglobin [pg]

MCHC mean corpuscular haemoglobin concentration [g/dL]

MCV mean corpuscular volume [fL]

RBC red blood cell count [10^6/uL]

RDW.SD red blood cell distribution width [fL]

RETC reticulocyte count [10^6/uL]

RETP reticulocyte percentage [%]

OFFscore OFF-score

ABPS Abnormal Blood Profile Score

These samples are assumed to represent normal population.

Note

One of the rows actually belongs to one of the authors of this package, who promises that he was not doped.

Source

Swiss Laboratory for Doping Analyses (LAD), with some calculations performed by the World Anti-Doping Agency (WADA).

---

blooddoping   *Blood samples from an athlete convicted of doping.*

Description

A dataset containing the result of the analysis of 13 blood samples, taken over a period of 5 years, of a female athlete who was convicted of doping on the basis of the Athlete Biological Passport.

Usage

```
blooddoping
```

Format

A data frame with 13 rows and 11 variables (including 10 haematological variables):

date the date of test

HCT haematocrit [%]

HGB haemoglobin [g/dL]

MCH mean corpuscular haemoglobin [pg]

MCHC mean corpuscular haemoglobin concentration [g/dL]

MCV mean corpuscular volume [fL]

RBC red blood cell count [10^6/uL]

RETC reticulocyte count [10^6/uL]

RETP reticulocyte percentage [%]

OFFscore OFF-score

ABPS Abnormal Blood Profile Score

In November 2012, the athlete was convicted of doping by the Court of arbitration for Sport, the evidence showing at least two occurences of doping, in summer 2009 (shortly before the test of 2 July 2009) and in June 2011, likely using an agent such as recombinant erythropoietin (rhEPO).

Doping is indicated in particular by high haemoglobin values associated with very low reticulocyte % and high OFF scores (a combination of these two variables).

Note

The data tables published in the original source contain several typos, as confirmed by the World Anti-Doping Agency (WADA). In particular, values for RETP in rows 8 and 10 were swapped and the MCHC value for row 7 was incorrect. This package's source code contains both the original data and the details of the corrections that were applied.

The ABPS values provided are very close (within <2%) to those obtained using the ABPS function, but some of them were likely calculated using different (older) versions of the ABPS code, which may explain some of these differences.

Source

https://jurisprudence.tas-cas.org/Shared%20Documents/2773.pdf

---

OFFscore *function OFFscore*

Description

The OFFscore function computes the value of the OFF-hr score (or OFF score), a combination of the haemoglobin level and the percentage of reticulocytes, used for detecting blood doping.

Usage

```
OFFscore(haemdata = NULL, HGB = NULL, RETP = NULL)
```

Arguments

haemdata	a vector or data frame containing (at least) the 2 haematological variables, either with the same names as the parameters below, or (not recommended) without names but in the same order as the parameters.
HGB	level of haemoglobin HGB [g/dL]
RETP	percentage of reticulocytes [%]

Details

The OFF-hr score is defined as 10*HGB [g/dL] - 60*sqrt(RETP [%]).

It is one of the parameters of the Athlete Biological Passport (ABP) program managed by the World Anti-Doping Agency (WADA), and is routinely used to identify athletes who use a substance prohibited by anti-doping rules (see e.g. https://jurisprudence.tas-cas.org/Shared%20Documents/4006.pdf).

The rationale for using this score for detecting blood doping is the following: if a manipulation (use of transfusion or of erythropoietic stimulating agents (ESA) such as recombinant human erythropoietin (rhEPO)) increases the number of circulating red cells (and thus increases the level of haemoglobin), the organism will react by stopping its own production of red blood cells. This negative feedback will be observed in the reduced percentage of reticulocytes (immature red blood cells). Such a combination of elevated HGB and reduced RETP, which will produce a high OFF score, is found neither naturally, nor as the consequence of a medical condition, and is thus indicative of doping.

The OFF score will pick up both withdrawal of blood (which induces a reduction in haemoglobin, a rise in reticulocytes, and thus a rise of RETP and a reduction of OFF score), and its re-infusion (HGB concentration increases, number of reticulocytes and RETP decrease, and OFF score increases).

Value

value of the OFF score.

Note

The values for HGB and RETP can be specified using either a data frame containing (at least) the 2 haematological variables, or using two named parameters, but not both at the same time.

The original OFF-hr score, as described by Gore et al., expects the haemoglobin level HGB to be specified in g/L. In order to be coherent with other functions in the package, this function assumes that HGB is specified in g/dL, and will multiply the value by 10. A warning will be emitted if units used seem wrong.

References

Gore, C.J., R. Parisotto, M.J. Ashenden, et al., Second-generation blood tests to detect erythropoietin abuse by athletes. Haematologica, 2003. 88(3): p. 333-44.

Examples

```
OFFscore(HGB=14.6, RETP=0.48)

data(blooddoping)
OFFscore(HGB=blooddoping$HGB, RETP=blooddoping$RETP)
```

Index