

1 **Manuscript title: Cardiovascular risk prediction: from classical statistical methods to**
2 **machine learning approaches**

3
4 **Running title: Computational approaches in cardiovascular risk prediction**

5
6 **AUTHORS:** Michela Sperti^{1*}, Marta Malavolta^{1*}, Federica Staunovo Polacco¹, Annalisa
7 Dellavalle¹, Rossella Ruggieri¹, Sara Bergia¹, Alice Fazio¹, Carmine Santoro¹, Marco A.
8 Deriu¹.

9 ¹ *Polito^{BIO}MedLab, Department of Mechanical and Aerospace Engineering, Politecnico di Torino,*
10 *Turin Italy;*

11 ** These authors contributed equally.*

12 *Corresponding Author (marco.deriu@polito.it).*

13
14 **Abstract**

15 Nowadays, cardiovascular risk prediction scores are commonly used in primary prevention
16 settings. Estimating the cardiovascular individual risk is of crucial importance for effective
17 patient management and optimal therapy identification, with relevant consequences on
18 secondary prevention settings. To reach this goal, a plethora of risk scores have been
19 developed in the past, most of them assuming that each cardiovascular risk factor is linearly
20 dependent on the outcome. However, the overall accuracy of these methods often remains
21 insufficient to solve the problem at hand. In this scenario, machine learning techniques have
22 repeatedly proved successful in improving cardiovascular risk predictions, being able to
23 capture the non-linearity present in the data. In this concern, we present a detailed discussion
24 concerning the application of classical versus machine learning-based cardiovascular risk
25 scores in the clinical setting. This review aims to give an overview of the current risk scores
26 based on classical statistical approaches and machine learning techniques applied to predict
27 the risk of several cardiovascular diseases, comparing them, discussing their similarities and
28 differences, and highlighting their main drawbacks to aid the physician having a more
29 critical understanding of these tools.

30
31 **Keywords:** Risk Prediction, Machine Learning, Traditional Risk, Artificial Intelligence,
32 Cardiovascular Diseases.

33

1 **Introduction**

2 Cardiovascular diseases (CVDs) are a group of disorders and conditions that affect the
3 structures and/or the function of the heart and blood vessels. They are the world's leading
4 cause of mortality. Due to the enlarged globalization and urbanization, it is noticed that the
5 risk factors for CVDs are rapidly increasing. About 17.7 million people die every year due
6 to CVDs and this number is estimated to be 31% of all global deaths. Nearly 75% of CVD
7 deaths occur in low-income middle-income countries. The main risk factors can be
8 classified into two categories: non-modifiable risk and modifiable risk factors; only the
9 latter can be modified with a healthier lifestyle with the aim of practising physical activity,
10 quitting smoking and focusing on treatments to lower blood pressure and change cholesterol
11 levels¹. The above-mentioned risks are shown in **Errore. L'origine riferimento non è stata**
12 **trovata.** The most relevant CVDs affecting the *heart* involve atrial fibrillation (AF),
13 myocardial infarction (MI), valvular heart disease, cardiomyopathies, and hypertensive
14 heart disease, which are related to hypertension. The most important CVDs involving the
15 *blood vessels*, which are known as vascular diseases, include aortic aneurysm,
16 atherosclerosis, and diseases that limit blood affluence to the brain (stroke) or the legs and
17 the arms.

18 In the 1950s, given the enormous incidence of CVDs, there was a need to develop predictive
19 models that could predict the onset of fatal events and be able to apply the most appropriate
20 therapy. The risk assessment and prevention at the primary stage is quicker to adopt and
21 implement than at the secondary stage of prevention and can potentially lead to a reduced
22 long-term cost upon the sanitary system. However, in clinical practice, it is easier to classify
23 a patient which presented a previous event (secondary prevention), and who is at high risk
24 rather than globally evaluate healthy patients based on identifiable risk factors (primary
25 prevention). In this scenario, it is of crucial importance to develop computational methods
26 able to detect as early as possible the presence of a disease and, more than that, the
27 probability of a patient to develop a certain pathological condition. The prevention of CVD
28 at the primary stage focuses on the detection of high-risk factors that are the major causes
29 of CVD. The key query to be resolved is how to define different risk levels and assess the
30 CVD risk.

31 The above-mentioned predictive models determine the probability of an individual
32 experiencing a CV event within a predefined period by assessing the entire risk factors
33 profile. Therefore, it is important to choose the most appropriate predictive model to avoid

1 misdiagnosis and expose patients to unnecessary treatments which would possibly lead to
2 side effects. The so-called global cardiovascular risk profile is therefore an indicator that,
3 given a set of risk factors, uses them to evaluate the probability of experiencing an acute
4 cardiovascular event. The idea of identifying people at risk based on factors considered
5 individually has been abandoned, but their overall assessment is taken into consideration.
6 For this reason, the absolute global risk is not the simple sum of the risk due to individual
7 factors.

8 To date, with the rapid development and widespread adoption of artificial intelligence (AI),
9 machine learning (ML), and deep learning (DL) techniques, they have become popular
10 methodologies also in the healthcare domain and have repeatedly proven successful in many
11 disease risk prediction scenarios²⁻⁴. The objective of employing ML to predict the
12 cardiovascular risk in the population is the improvement of current risk scores performances
13 and the exploitation of the hidden information present in the big amount of data available.
14 Traditional methods cannot manage large datasets, since they are limited to a certain number
15 of variables⁵.

16 AI is the computer science subfield involved in systems' creation that performs tasks that
17 would require human intelligence. It also includes ML and DL. ML focuses on approaches
18 that allow machines to learn from data without being explicitly programmed; it tends to
19 minimize errors and maximize the probability that their predictions are reliable. The three
20 main ML classes are supervised learning, unsupervised learning, and reinforcement
21 learning. DL incorporates computational models and algorithms that mimic biological
22 neural networks in the brain⁶.

23 ML methods may be much more valuable than traditional methods to predict the probability
24 that an event will occur without being associated with specific risk factors. It is not always
25 possible to make a diagnosis with only the data of standard clinical trials; ML helps to make
26 a diagnosis quickly without wasting time for additional research and at the same time serves
27 to expand and increase the efficiency of treatments of patients with CVD. Thanks to the
28 evolution of technology, patients can monitor their state of health through mobile devices,
29 and in case they can realize in real-time a worsening of their condition. The main objective
30 of ML is to incorporate data from different sources (clinical measurements, biological data,
31 experimental results, environmental information).

1 The present review aims at summarizing current risk scores based on classical statistical
2 approaches and ML techniques and discussing their main clinical applications in the context
3 of cardiovascular risk prediction (Figure 2).

4 **1 Classical Statistical Methods**

5 In the following sections, a detailed description of the most famous risk scores based on
6 traditional statistical methods is provided with the most relevant clinical applications:
7 Framingham Risk Score (FRS), Systematic COronary Risk Evaluation (SCORE), QRISK,
8 PROspective CARdiovascular Münster (PROCAM), Reynolds Risk Score (RRS) and
9 CUORE.

10 **1.1 Framingham Risk Score (FRS)**

11 The Framingham Heart Study (FHS) is the longest-running cardiovascular epidemiological
12 cohort study in the USA, which established that age, sex, lifestyle, hypertension, total
13 cholesterol, high-density lipoprotein (HDL) cholesterol, systolic blood pressure,
14 antihypertensive medication use, dyslipidemia, smoking, and diabetes are risk factors for
15 CVDs, and urged the development of efficient treatments for these conditions. The result
16 was a greatly diminished incidence of CVDs⁷. The FHS defines CVD as a composite of
17 coronary heart diseases (CHD: angina, coronary death, coronary insufficiency, MI),
18 cerebrovascular events, heart failure, and peripheral artery disease⁸. The FRS is a
19 multivariable risk function used to point out the possibility of developing CVDs. FRS
20 predicts risks of CVD by using Cox proportional risks regression⁹, which is a technique that
21 allows examining the correlation between the incidence of a clinical outcome (such as MI)
22 and a risk factor (such as smoking). The risk factors taken into account were age, systolic
23 blood pressure, antihypertensive medication use, total cholesterol, HDL cholesterol,
24 smoking, and diabetes. Other variables (diastolic blood pressure, body mass index,
25 triglycerides) also were considered, but they were not statistically relevant. Employing Cox
26 proportional risks regression, a CVD risk prediction function was estimated, to evaluate the
27 10-year absolute CVD risk. All the continuous variables were naturally logarithmically
28 converted to improve the discrimination and calibration of the models. The developed
29 function is the following:

$$30 \quad FRS = 1 - S^0(t) \exp(\sum_{i=1}^p b_i X_i - \sum_{i=1}^p b_i \bar{X}_i) \quad [1]$$

31 where $S_0(t)$ is the baseline survival at follow-up time t (here $t=10$ years), b_i the estimated
32 regression coefficient (log hazard ratio, the values of which are shown in Table I), X_i the

1 log-transformed value of the i^{th} risk factor, \bar{X}_i the corresponding mean, and p is the number
2 of risk factors. However, this study shows some limitations: FRS is not applicable in Europe
3 because it is based on parameters referred to the North-American population. For this
4 reason, another method for risk calculation was developed in the past (namely SCORE
5 algorithm), which is applicable in Europe. Furthermore, FRS doesn't consider some
6 important risk factors, like CVD family history and lifestyle.

7

1

Table I. Regression Coefficients and Hazard Ratios¹². SBP = systolic blood pressure.

Variable	Women		Men	
	β	Hazard Ratio	β	Hazard Ratio
Log of age	2,33	10,27	3,06	21,35
Log of total cholesterol	1,21	3,35	1,12	3,08
Log of HDL cholesterol	-0,71	0,49	-0,93	0,39
Log of SBP if not treated	2,76	15,82	1,93	6,91
Log of SBP if treated	2,82	16,82	2,00	7,38
Smoking	0,53	1,70	0,65	1,92
Diabetes	0,69	2,00	0,57	1,78

2

3 The FHS was helpful to *evaluate the impact of AF* on the risk of death, after correcting for
4 coexisting cardiac diseases and risk factors in a population-based sample, using the pooled
5 logistic regression analysis¹⁰. The clinical risk factors for AF are advancing age, diabetes,
6 hypertension, congestive heart failure, rheumatic and nonrheumatic valve disease, and MI.
7 The pooled logistic regression analysis is a Cox regression analysis that depends on time,
8 so it was possible to see how the risk factors changed over time with the clinical variables.
9 About two subjects without AF were matched to each AF subject by gender, age, and date
10 of diagnosis of AF. The differences in survival between AF and matched non-AF
11 participants were analysed to estimate survival and produce mortality curves. It has been
12 noticed that cardiovascular disease risk factors were more frequent in subjects with AF than
13 subjects without AF. In addition, the mortality of men and women with AF was much higher
14 than for the non-AF subjects. At 10 years of follow-up, in people aged between 55 and 74
15 years, 57.6% of women with AF had died compared with 20.9% of women without AF, and
16 61.5% of men with AF had died, whereas 30.0% of men without AF had died. Furthermore,
17 in men with AF and aged between 55 and 64 years, 12.6 years of survival have been
18 estimated as against 18.1 years in men without AF, and in women 12.1 years of survival
19 were estimated compared with 21.3 years in women without AF. So, AF is significantly
20 associated with an increased risk of death.

21 A recent application of FRS is related to the *COVID-19*. It was noticed that patients admitted
22 to hospitals with existing CVDs, experienced worse outcomes compared to CVD-free
23 patients. In fact, in a sample of 700 hospitalizations for COVID-19, unfavourable levels of
24 all individual CVD risk factors were related to an almost doubled risk of subsequent
25 COVID-19¹¹. For people in the highest risk groups, the FRS offered some predictive utility

1 helpful for the identification of at-risk groups, which now are the first groups to be given
2 the vaccine.

3 **1.2 Systematic COronary Risk Evaluation (SCORE)**

4 The SCORE is a method developed in 2003 from the Framingham Heart Study¹²; this
5 method aims to evaluate the 10-year risk to run into a first fatal cardiovascular event. As the
6 Framingham method was not suitable for the European population, a researchers' group
7 developed the so-called SCORE; it also allows the division between high CV risk countries
8 and low CV risk countries and gives more reliable results in people aged between 40 and 65
9 years. SCORE allows to calculate (through the use of specific charts) the vascular age of a
10 person, that is to say, the age of the vascular system concerning exposure to certain risk
11 factors, including total cholesterol, systolic blood pressure, gender, age, and smoking status;
12 without making any distinction between diabetic patients or non-diabetic patients.

13 The vascular age may not coincide with the biological age of the person, thus exceeding the
14 limit for the calculation of cardiovascular risk based only on the age of the person. It also
15 demonstrates that it's not always true that elderly people are more exposed to cardiovascular
16 risks: it depends on how many years of exposure to risk factors they have. The concept of
17 vascular age also improves communication with patients, making them more aware of the
18 risk they run into if they do not change their lifestyle. For example, a 50-year-old woman
19 who suffers from high blood pressure (180 mmHg), has a high level of cholesterol in her
20 blood, and smokes risks as much as a 65-year-old woman who has normal blood pressure
21 (140 mmHg), does not smoke, and has a lower cholesterol level. These charts differ if
22 applied to high CV risk countries (Finland, Russia, Norway, British region, Scotland,
23 Denmark, Sweden, and Germany) or low CV risk countries (Belgium, Italy, France, and
24 Spain)¹³.

25 The most important application using the SCORE is a *comparison between FRS and SCORE*
26 *in the Australian population*. A specific study was conducted in Australia in which the
27 results obtained with the FRS method were compared with the SCORE method for both
28 high-risk and low-risk countries, as they have similar endpoints, both consider the same risk
29 factors and variables, and both models predict the 10-year CVD death risk¹⁴. The
30 participants were 4487 women aged 30-74 years, from Australia and free of cardiovascular
31 events. For each patient, the mortality risk was calculated using the FRS and the two SCORE
32 risk models, and these were compared. The accuracy of the CVD risk score models was
33 evaluated using both discrimination and calibration. During the 10-year follow up there

1 were 152 deaths, of which 28 were caused by cardiovascular diseases, 3.4% of the sample
2 (**Errore. L'origine riferimento non è stata trovata.**). All two methods underestimated the
3 death risk in women below 50 years of age and overestimated the risk in those over the age
4 of 50. The FRS model and SCORE risk chart for low-risk regions predicted similar risk
5 levels in the age categories, but at the same time, the SCORE high-risk model predicted
6 higher risk levels compared to the other two models.

7 Another application that involves an updated version of SCORE, is called *SCORE OP*. It
8 was developed in 2015 to predict the risk of fatal CV events specifically for patients over
9 the age of 65 in Germany¹⁵. The group of people consisted of 1657 participants aged 70 and
10 over who had never had cardiovascular problems before, and during the 5 -year follow-up,
11 118 cardiovascular deaths occurred due to fatal MI, fatal coronary heart disease, sudden
12 cardiac death, death due to other cardiac diseases such as heart failure. The risk factors
13 considered were age, gender, systolic blood pressure, HDL and total cholesterol, smoking
14 status, diabetes mellitus status, and self-reported use of antihypertensive medication. The
15 results show an overestimation of the risk and specifically, SCORE OP predicted higher
16 risks compared to SCORE in female individuals aged >75 and in male individuals aged >
17 78 (**Errore. L'origine riferimento non è stata trovata. Errore. L'origine riferimento non**
18 **è stata trovata.**). Although there has been an overestimation of risk, SCORE OP has given
19 a result more discriminatory than SCORE.

20 A particular implementation of the SCORE is the *IberScore study* in Spain. The goal of this
21 method is the computation of the cardiovascular risk in young, healthy, and working
22 people¹⁶. In contrast to the original score method, this model separates women from men in
23 two different categories (it does not consider gender as a mere risk factor). The number of
24 people examined by the study was 624389 workers aged 16 to 65, with no cardiovascular
25 history. Among the risk factors, however, there are smoking status, total cholesterol and
26 HDL, systolic blood pressure (SBP), and in addition blood glucose, obesity, and another
27 predictor that summarised a personal history of dyslipidemia, diabetes, hypertension, and/or
28 adherence to a prescribed heart-healthy diet that SCORE did not consider before. They
29 observed 3741 cardiovascular events during the 10-year follow-up (80.5% non-fatal and
30 19.5% fatal) and the incidence rate was four times greater for men. The results suggest that
31 under 55 years, SCORE underestimated the risk, while Iberscore showed higher risk
32 predictions. Over 55 years, on the contrary, SCORE overestimated the risk, while Iberscore
33 gave more reliable predictions.

1 1.3 QRISK

2 QRISK is a prediction algorithm for CVD used in the United Kingdom according to the
3 National Institute for Health and Care Excellence (NICE)¹⁷. This model provides an
4 estimate of the 10-year risk of experiencing a CVD event, and it is widely used in the context
5 of primary care. Patients with a QRISK score below 20% are considered to be at low risk
6 of developing CVD, whereas patients with a QRISK score above 20% are considered to be
7 at high risk and therefore should be monitored and possibly initiate supportive therapies.
8 The QRISK model was developed from a research database called QRESEARCH which
9 collects sufficient routine data from the anonymised health records of over 35 million
10 patients derived from general practices using the EMIS clinical computer system
11 (<https://www.emishealth.com/>). The practices are distributed throughout the UK. Historical
12 records (which also include patients who may have died or left) extend back to 1989 making
13 it one of the richest general practice databases in the world (<https://www.qresearch.org/>).
14 The research uses two-thirds of the database to derive and model the algorithm and the
15 remaining one-third to control and validate the sample. QRISK uses the following risk
16 factors: age, gender, smoking habit, SBP, body mass index (BMI), the ratio of total serious
17 cholesterol to HDL cholesterol or lipoprotein levels, ethnicity, deprivation measures, family
18 history of CVD, AF, RA, chronic kidney disease, diabetes mellitus and use of at least one
19 blood pressure treatment such as thiazide antihypertensive blocker, calcium channel blocker
20 or angiotensin-converting enzyme inhibitor.

21 There are several versions in the QRISK model, which differ in some risk factors. For
22 example, QRISK2 includes self-assigned ethnicity, type 2 diabetes, kidney disease, RA, and
23 AF, which are not considered in QRISK1. Furthermore, QRISK2 relates age with other
24 factors such as BMI, family history, or smoking habit¹⁸. QRISK model uses a logarithmic
25 transformation for the 'age' variable and linear terms for the rest of the variables according
26 to the fractional polynomial analysis.

27 An interesting aspect is represented by the inclusion of deprivation and ethnicity within the
28 risk factors of QRISK. This aspect could help prevent and reduce the widening of health
29 inequalities. This feature distinguishes the QRISK score from other predictive methods such
30 as the Framingham Risk Score and is very useful for correctly estimating cardiovascular
31 risk. For example, a conducted study showed that 9% of the total sample was reclassified
32 from low to high risk or vice versa by using QRISK versus FRS and in this article, it is
33 emphasized that FRS algorithm tends to underestimate cardiovascular risk in disadvantaged

1 areas and overestimate cardiovascular risk in affluent areas. An external validation study
2 was done by an Oxford University independent team which used a dataset external to the
3 QRESEARCH database showing that QRISK is better calibrated on the population of the
4 United Kingdom compared to other predictive models¹⁹. QRISK resulted better calibrated
5 than FRS. This is because FRS does not include risk factors such as BMI, family history of
6 cardiovascular disease, antihypertensive treatment, and social deprivation.

7 In a 2017 study, it is shown a correlation between the percentage of cardiovascular risk
8 recorded by QRISK and the use of lipid-lowering therapy treatments such as statins. Statins
9 are a group of molecules that can lower blood cholesterol levels and consequently reduce
10 the risk of heart attacks and strokes (up to 25%). A higher than 10% QRISK score suggests
11 the use of statin therapies for prevention. The data support this thesis by indicating that the
12 majority of patients treated with statins did not register a QRISK score and the majority of
13 patients at high risk according to the QRISK, were not initiated with statin therapy²⁰.

14 Another interesting aspect is the implementation of a model called the *QRISK life model* to
15 estimate the individual risk of cardiovascular disease during life. The algorithm is based on
16 the QRISK2 model and uses data routinely collected by UK general practice. Thanks to this
17 algorithm it is possible to define centile values that can be used as thresholds to classify
18 further high-risk patients. If the analysed patient has a value higher than the 90th centile or
19 lifetime risk of CVD higher than 50%, they will be at high risk during their lifetime²¹. For
20 example, a 54-year-old white woman who does not smoke, obese, coming from an affluent
21 area, with a first degree relative with CHD, SBP of 150 mmHg, and HDL cholesterol ratio
22 of 5.3, has a risk of cardiovascular disease at 10 years of 14%, therefore below the
23 intervention threshold at 10 years of 20%. However, her lifetime risk of cardiovascular
24 disease is 62%, which places her above the 95th centile for life risk. If we compare the same
25 person with better risk factors, namely BMI 28, systolic blood pressure 128 mmHg, total
26 HDL cholesterol ratio 4, they have a risk of 49% throughout life, and risk under 10% at 10
27 years. With the current risk factors, the expected age for the occurrence of a cardiovascular
28 event is 81 years, but with a reduction in risk factors, it shifts to 95 years.

29 Overall, patients with a high lifetime risk are more likely to be male, from non-white ethnic
30 groups (particularly South Asia), current smokers, and with a positive family history of
31 CHD compared to the ones with a 10-year high risk. The QRISK life model could be used
32 on an individual level to communicate cardiovascular risk to patients, particularly young
33 patients with positive family histories whose early lifestyle changes or drug interventions

1 could lead to significant improvements. It remains questionable whether therapies should
 2 be based on 10-year risk or lifetime risk. On one hand, early lifestyle interventions may be
 3 helpful but on the other hand, medical operations carry risks as soon as they are initiated.
 4 Furthermore, the main disadvantage of this algorithm is that the values and results found
 5 are only applicable to patients from the UK, since it has been developed on the British
 6 population of QRESEARCH database (<https://www.qresearch.org/>).

7 **1.4 PROspective Cardiovascular Münster (PROCAM)**

8 Among the various models which calculate this risk profile, many of these (e.g., FRS) do
 9 not consider many important features, like LDL cholesterol or family history of CHD. This
 10 is done by the PROCAM score, based on a study carried out in Germany²², which involved
 11 for its development 5389 men with ages between 35 and 65. By the 10 year follow-up, 325
 12 acute coronary events resulted. To predict the CHD, the “simple” clinical judgment is not
 13 sufficient, because it involves too much subjectivity; thus, statistical methods have been
 14 implemented by PROCAM. These give a more realistic percentage of risk. In that case, the
 15 method consisted of a Cox proportional hazards model²³, which has many advantages, like
 16 logistic regression. It used 8 independent risk variables, listed in order of importance: age,
 17 LDL cholesterol, smoking (yes or no), HDL cholesterol, SBP, family history of premature
 18 MI (yes or no), diabetes mellitus (yes or no), and triglycerides. Of the 57 clinical features
 19 measured in the PROCAM study, 8 were used to construct the risk algorithm because they
 20 were found to be independently predictive of event risk. These variables, together with the
 21 β -coefficients of the Cox model, the hazard ratios, and the 95% confidence intervals, are
 22 shown in Table II²². The table also shows the mean value of each variable in the total study
 23 population.

24
 25 *Table II. Variables used in the construction of the Cox proportional hazards model. The P-value is <0,001*
 26 *for all variables.*

Variables	β	Hazard Ratio
Age, y	0,103	1,108
LDL cholesterol, mg/dl	0,013	1,013
Smoking, %	0,658	1,931
HDL cholesterol, mg/dl	-0,032	0,968
SBP, mmHg	0,010	1,010
Family History of MI, %	0,382	1,465

Diabetes mellitus, %	0,399	1,491
Triglycerides, mg/dl	0,317	1,373

1
2
3
4
5
6
7
8
9
10
11
12

With the Cox model, only the relative risk can be calculated. Relative risk is the ratio of the proportion of patients who present the event in the risk group and the proportion of patients experiencing the event in the group not exposed to the risk factor. Consequently, to transform the relative risk into an absolute one, the Kaplan-Meier statistic was used, which permits the construction of the survival curves (i.e., the graph of the relationship between the cumulative probability of surviving and the observation time) and to measure the observed risk. To increase the easiness and usefulness of the risk algorithm in clinical practice, a simple scoring scheme was created: each risk factor was divided into classes, as can be seen in Table III²².

Table III. Coefficients of PROCAM Scoring Scheme.

Age, y		LDL, mg/dL		SBP, mmHg		Triglycerides, mg/dL	
<i>Range</i>	<i>Score</i>	<i>Range</i>	<i>Score</i>	<i>Range</i>	<i>Score</i>	<i>Range</i>	<i>Score</i>
35-39	0	<100	0	<120	0	<100	0
40-44	6	100-129	5	120-129	2	100-149	2
45-49	11	130-159	10	130-139	3	150-199	3
50-54	16	160-189	14	140-159	5	≥200	4
55-59	21	≥190	20	≥160	8		
60-65	26						
Diabetes mellitus		Smoker		HDL, mg/dL		MI in family history	
<i>Range</i>	<i>Score</i>	<i>Range</i>	<i>Score</i>	<i>Range</i>	<i>Score</i>	<i>Range</i>	<i>Score</i>
No	0	No	0	<35	11	No	0
Yes	6	Yes	8	35-44	8	Yes	4
				45-54	5		
				≥55	0		

13
14
15
16
17
18

Then the global risk was calculated by use of the Cox model in combination with the survival curves. After that, a regression equation was calculated between the logarithm of global risk and the classes of each risk factor. The coefficients obtained in this way were then standardized and rounded to the integer. Table IV shows the risk of a coronary event associated with each score, calculated using the full PROCAM algorithm²².

1

2

Table IV. 10 years risk of acute coronary events associated with each PROCAM score²².

Points	Risk (%)	Points	Risk (%)	Points	Risk (%)	Points	Risk (%)
≤20	<1.0	36	4.2	25	1.6	51	16.8
21	1.1	37	4.8	26	1.7	52	17.5
22	1.2	38	5.1	27	1.8	53	19.6
23	1.3	39	5.7	28	1.9	54	21.7
24	1.4	40	6.1	29	2.3	55	22.2
30	2.4	41	7.0	45	10.2	56	23.8
31	2.8	42	7.4	46	10.5	57	25.1
32	2.9	43	8.0	47	10.7	58	28.0
33	3.3	44	8.8	48	12.8	59	29.4
34	3.5	45	10.2	49	13.2	≥60	≥30.0
35	4.0	46	10.5	50	15.5		

3

4 The accuracy of the simple point scoring system was similar to coronary event prediction
5 using continuous variables, with the important difference that the use of the scoring scheme
6 gives more accurate targeting of lipid-lowering therapy and strong cost savings. Indeed,
7 considering the so-called ROC (receiver-operating characteristics) curve, that draws the
8 ratio between true positive rate and false positive rate, whose ideal value is 1 (100%), it
9 shows an area under the curve of 82.4%, compared with 82.9% for the Cox model (with
10 continuous variables).

11 Trying to use the FRS to determine the relative risk of an acute coronary event and
12 evaluating it through ROC curve analysis, the area under the curve was 77.8%, so it was
13 significantly less than that obtained with the PROCAM models. There are, however,
14 important differences between the PROCAM and the FHS data sets. For example, FRS does
15 not include information on family history of CHD, triglycerides, or LDL cholesterol. The
16 Framingham risk prediction function systematically overestimated risk in the PROCAM
17 cohort. The PROCAM score performs better (than FRS) with the data set in which it was
18 developed and optimized. This is a classical characteristic of the classical statistical methods
19 (which will be improved by using ML techniques, described later). In fact, more real
20 validation of both the Framingham and PROCAM scoring systems would need their
21 application to a third independent data set.

1 An example of PROCAM risk score application is represented by a systematic review of
2 published research in which the score was used together with FRS to test if bariatric surgery
3 reduces long-term CV events²⁴. CV risk models were applied in a validation cohort
4 previously published. The validation cohort was a historical cohort of 197 patients who
5 underwent Roux-en-Y gastric bypass and 163 control patients (with nonoperative
6 treatment), identified through the Rochester Epidemiology Project. FRS and PROCAM risk
7 scores were applied to calculate 10-year CV risk. In the validation cohort, the absolute 10-
8 year FRS for CV events was lower at follow-up in the bariatric surgery group (7.0% to
9 3.5%, $p<0.001$) compared with controls (7.1% to 6.5%, $p=0.13$), with an absolute difference
10 in risk reduction between the two groups of 3% ($p<0.001$). PROCAM risk in the bariatric
11 surgery group decreased from 4.1% to 2.0% ($p<0.001$), whereas the control group exhibited
12 only a modest decrease (4.4% to 3.8%, $p=0.08$). Comparing the average data from the
13 validation study with the Roux-en-Y group, it was noticed that the tendency of CV risk was
14 similar. Thus, bariatric surgery offers significant reductions in predicted CV risk.
15 Furthermore, some studies demonstrated a mortality benefit after bariatric surgery. This is
16 because obesity not only increases cardiovascular risk, but also the risk of developing
17 respiratory diseases, certain cancers, and more. In a study with a 7.1-year mean follow-up
18 period, long-term mortality from any cause was decreased by 40% compared with controls,
19 and mortality from CV disease was decreased by 56%²⁵. In conclusion, risk models could
20 be applied to bariatric surgery patients to inform them of their future risk for CV events and
21 all-cause death.

22 **1.5 Reynolds Risk Score (RRS)**

23 RRS for the prevention of CVD has been developed in 2007 for women and men. The
24 difference between RRS and the elder Risk Scores such as FRS concerns the features
25 involved in evaluating the risk classification of the various patients. For example, while in
26 FRS the main features used for the classification are age, smoking, blood pressure,
27 antihypertensive treatment, diabetes mellitus, and total and high-density lipoprotein
28 cholesterol, RRS adds some features as high-sensitivity C-reactive protein, family history
29 of premature MI, and haemoglobin A1c (among diabetics only). It does not predict CVD in
30 particular but it calculates a composite CVD outcome that may refer to MI incident,
31 ischemic stroke, coronary occlusion, and death due to CVD in general²⁶. A detailed example
32 of statistical analysis of population, creation, and validation of RRS algorithm has been
33 described by the Women's Health Initiative (WHI). The algorithm has been validated on a

1 set, while one-third had the role of the validation set. The first part of the process consisted
 2 of the determination of the model using the Cox reduction method using two-thirds of
 3 collected data, and the second part is centred on the validation of the model using one-third
 4 of the total population²³. The extendibility of RRS is based on the determination on the
 5 Bayes Information Criterion (BIC), a sort of regularization in which lower values indicate
 6 a better fitting and the increasing of variable induces a penalty. BIC usually describes the
 7 goodness of the fitting of a statistical method to sample data for given values of unknown
 8 parameters. The coefficients for RRS are shown in Table V.

9
 10 **Table V.** This table shows the coefficient β and standards errors (SE) for RRS. The P-value is $<0,001$ for all
 11 variables. Abbreviations: HbA_{1c} = hemoglobin A1c; SBP = Systolic blood pressure; HDL-C = high-density
 12 lipoprotein cholesterol; hsCRP = high- sensitivity C-reactive protein; MI = myocardial infarction²⁷.

Variables	β (SE)	χ^2
Age	0,080 (0,006)	193,5
HbA _{1c} % with diabetes	0,134 (0,017)	62,3
Current smoking	0,818 (0,109)	55,9
ln(SBP)	3,137 (0,423)	55,1
ln(HDL-C)	-1,172 (0,172)	46,2
ln(Total Cholesterol)	1,382 (0,239)	33,3
ln(hsCRP)	0,180 (0,043)	17,5
Parental history of MI < Age 60 y	0,438 (0,118)	13,7
Age	0,080 (0,006)	193,5

13
 14 Coefficient β is also called “regression coefficient” and is used to calculate how well the
 15 curve that predicts the outcomes is near to the sample that we analyse. The coefficient χ^2
 16 indicates the “standard error” and estimates the accuracy of the algorithm in predicting
 17 outcomes. Finally, p-value also called “probability value”, evaluates the probability to
 18 obtain outcomes similar to results obtained during the testing of the algorithm. In other
 19 words, p-value helps to understand if the difference between obtained and supposed
 20 outcomes is due to fortuity, low levels of p-value indicate that can be a very significant
 21 difference that cannot be attributed to randomness in sampling. Moreover, the BIC
 22 coefficient for RRS reaches a very low value underlining a good predicting ability with
 23 clinical simplifications. RRS could be confronted with the most used risk scores, such as

1 FRS or ATP-III prediction model. FRS and ATP-III present very different values of BIC if
2 reported to RRS that shows a loss of predicting ability with clinical simplification.

3 RRS has been confronted with other algorithms by WHI in a scientific essay led by Dr.
4 Nancy Cook²⁶. The group of researchers analysed the predictive curve of the main used risk
5 scores, the results have been organized in tabs where the main features observed are:
6 calibration and accuracy. The first is the ability of the curve to fit the parameters and
7 discriminate the various subjects basing on their parameters, and the second is directly
8 linked to the ability to predict in the right way the outcome, in other words, it is based on
9 the right correspondence between supposed and obtained outcome. The average risk
10 obtained from the different risk scores was 3.8%, 4.6%, and 10,9% respectively for ATP-
11 III, RRS, and FRS, highlighting an intermediate value of RRS between the two most used
12 risk scores. In general, RRS seems to be more calibrated in the determination of CVD since
13 the other way to calculate CVD risk score tends to overestimate or underestimate the risk
14 according to WHI studies. After 10 years from the beginning of the screening, WHI
15 collected all data from patients controlling the appearance of CVDs among the samples.
16 Data based on the course of cardiological clinic painting were compared with the prediction
17 curves typical of the various risk score to understand the accuracy and the calibration of the
18 most used risk score and value the most suitable on a particular sample.

19 RRS according to the studies of Cook, Nancy and co-workers, provides a very effective way
20 to calculate the patient's risk to occur in a cardiovascular event in the following 10 years
21 since the determination of the risk score²⁶. Using RRS instead of elder risk scores (such as
22 FRS, SCORE, and ATP-III) guarantees a better calibration and accuracy of the prediction
23 curve, and this important improvement makes it possible to create a more accurate diagnosis
24 minimizing the risk to underestimate or overestimate the outcomes. Moreover, RRS
25 includes the quantity of C-reactive protein, haemoglobin A1c and cardiovascular history
26 among the relevant features used to evaluate the risk score. The quantity of reactive C-
27 protein increases during inflammatory events, and in this way RRS becomes useful in
28 considering systemic inflammatory processes in the prediction phase²⁸. High levels of
29 reactive C-protein are strongly linked to CVD as ischemic stroke and ictus since blood
30 vessels seem to be very sensitive to PCR and a great amount of it can cause damns or
31 lacerations of arterial and venous tissues. The addition of PCR and CVD history among risk
32 factors makes RRS an effective prediction method in patients with inflammatory diseases
33 and helps physicians evaluating the possibility to begin statin therapy. Statins are a group

1 of drugs used to reduce the number of lipids in the organism, including cholesterol. Patients
2 with type 2 diabetes mellitus (DM) are at an increased risk for atherosclerosis and the great
3 quantity of glucose in a patient's blood can compromise the endurance of the blood vessels.
4 However, RRS can monitor the level of PCR in blood, which is one of the most important
5 markers to predict the onset of type 2 DM, and metabolic syndromes are often preceded by
6 weak but numerous inflammatory events²⁸. Hemoglobin A1c is considered only in diabetics
7 patients, so diabetes results to be another important marker for the prediction of CVD. So
8 this method can be used in presence of patients with diseases as diabetes or systemic
9 inflammatory events and provides an accurate risk prediction involving factors brought by
10 other diseases²⁸. Despite the great performances of RRS, Buring & Cook highlights the
11 possibility to generate an even more accurate and calibrated risk score using the Cox
12 reduction method and minimizing BIC^{23,27}.

13 **1.6 CUORE**

14 In 1998 the CUORE project - Epidemiology and Prevention of Cerebral and CVDs - was
15 launched, under the coordination of the *Istituto Superiore di Sanità* (ISS), to estimate the
16 probability of a first cardiovascular event occurs within 10 years of assessment²⁹. The idea
17 behind this is to monitor CV risk factors, lifestyles, and diseases, to picture the situation of
18 the population's health and to identify areas with a priority of intervention, and to develop
19 preventive actions³⁰. It is known that the Framingham method⁹, which resulted very suitable
20 for the American population, cannot fit with the same accuracy populations registering a
21 lower rate of CHD, such as the European ones³⁰. Indeed, in this case, the Framingham
22 method-based risk charts tend to overestimate the absolute risk. More reliable would be
23 applying an algorithm of risk validated on the population of reference. In this scenario, the
24 CUORE risk score, developed during the CUORE project, answers the need for a better
25 fitting tool of evaluation on the Italian population, and others with low CHD rate³¹. It is
26 based on the mathematical equation [3] to do an appraisal of the probability of a first major
27 CV event³².

$$28 \quad CUORE = 1 - S(t)e^{[\sum_{i=1}^p \beta_i X_i - G(\mu)]} \quad [3]$$

29 where $S(t)$ is the survival at 10 years, β_i are the risk factor coefficients, performed using the
30 Cox proportional hazards model, separately for men and women²³ (Table VI), x_i is the i^{th}
31 factor of risk, $G(\mu)$ is the linear combination of the mean of the factors, and p is the number
32 of the considered factors of risk.

The project's official website <http://www.cuore.iss.it/> offers the possibility to consult collected data, calculated risk, and statistics, to have a global picture of the population's health situation. Moreover, it offers different options to calculate and evaluate the CV risk.

Table VI. Cox proportional hazard model for the CV assessment in 10 years³¹.

Variable	Women		Men	
	β	Hazard Ratio	β	Hazard Ratio
<i>Age</i>	0,07	2,05	0,08	2,12
<i>SBP (mmHg)</i>	0,01	1,29	0,02	1,39
<i>Total Cholesterol (mg/dl)</i>	0,01	1,32	0,01	1,24
<i>Hdl Cholesterol (mg/dl)</i>	-0,01	1,17	-0,02	1,28
<i>Smoking Status</i>	0,48	1,62	0,79	2,20
<i>Diabetes</i>	0,40	1,49	0,56	1,74
<i>Antihypertensive Treatment</i>	0,34	1,41	0,47	1,60

The first and most precise one is the calculation of an Individual Score, obtained either from the downloadable executive software, practical and easy to use in a medical environment, or filling in the online survey. This score tends to be more precise since it considers continuous values for some factors of risk and includes more variables. There are eight factors of risk: sex, age, diabetes, smoking habit, SBP, total and HDL cholesterol, and hypertensive treatment. It presents, though, some limitations: measurements must follow a standard methodology; age must be between 35 and 69 for both sexes and subjects must be without previous CV events; it is not suitable for pregnant women; it cannot be applied to outliers of the risk factors: SBP higher than 200 mmHg or lower than 90 mmHg, total cholesterol higher than 320 mg/dl or lower than 130 mg/dl, HDL cholesterol higher than 100mg/dl or lower than 20 mg/dl; clinical exams of glycemia and cholesterol should be not older than three months.

The other possible option is to estimate the risk score with the Cards of Risk. These are absolute global risk classes calculated by categories of risk factors, considering six factors of risk: sex, age, diabetes, smoking habit, SBP, total cholesterol. This score is less precise since it considers fewer factors, therefore it is possible to obtain different results with the two methods. Similar to the previous one, it presents some limitations. All the previous prescriptions are valid, but, in this case, they apply only to people aged 40-69 years old. In any case, it is recommended to periodically assess the CV risk, according to the risk range

1 of belonging. Furthermore, the project provides periodical surveys on the Italian population
2 of all the Italian territory, to keep the algorithm updated to the actual features of the
3 population, and took part in many projects for CVD prevention, and has an active role to
4 this extent.

5 Lately, a *comparative study between CUORE and Framingham algorithms* was carried out
6 to study their reliability in an Italian population³². There have been considered 996 subjects
7 of both sexes, patients of General Practitioners working in the Province of Rome. R2 value
8 has been used to evaluate the goodness of the regression model. Both the methods aim to
9 estimate the risk of CHD, but they differ in the variable to consider: Framingham considers
10 angina pectoris, recognized and unrecognized MI, coronary insufficiency (unstable angina),
11 and CHD deaths too, while CUORE can be applied to subjects who haven't had any
12 previous cardiovascular event, and the risk factors are obtained from the use of standardized
13 methodologies. Differently from the majority of the published studies, it emerged that FRS
14 is more reliable than CUORE on this specific Italian population (CVD risk is less
15 overestimated by FRS with respect to CUORE). The reason for such results can be maybe
16 found in the fact that all these algorithms tend to fit some populations better than others, and
17 maybe the features of the considered patients differ from the ones from other studies.
18 However, no unambiguous indication of the best algorithm to estimate the individual
19 cardiovascular risk is still available. On the other hand, the CUORE algorithm was revealed
20 to be more objective than the Framingham when it comes to the presence of other medical
21 conditions, such as psoriasis or severe mental illnesses (SMI).

22 CUORE was also employed in patients affected by rheumatic diseases. Research from
23 2013²⁹, aiming to investigate the relation between CV risk and psoriasis features in a real-
24 world setting, highlighted how the Framingham score on the Italian population was not
25 corresponding to the real evidence, while the CUORE resulted closer to it, proving the
26 connection of the two conditions. A few years later, an Italian bicentric study³³ explored the
27 performances of different algorithms in the prediction of cardiovascular events in patients
28 with rheumatoid arthritis (RA), another chronic inflammatory autoimmune disease that
29 electively affects the joints. Two Italian cohorts were used, and the discriminatory ability
30 for CV risk prediction was evaluated by analysing the area under the ROC curves, while the
31 calibration between predicted and observed events were assessed by Hosmer-Lemeshow
32 (HL) tests and sensitivity and specificity calculated for low to intermediate and intermediate
33 to high-risk cut-offs. CV events considered were: heart failure, sudden cardiac death, TIA,

1 CAD (stable and unstable angina pectoris), MI, PAD, and CVA. In this case, FRS, SCORE,
2 CUORE, Reynold's Risk Score (RRS), and QRISK2 tended to assess the CV risk lower
3 than what was possible to observe in reality in patients with RA. Since this category of
4 patients seems to be more exposed to CV problems, recently the EULAR (European League
5 Against Rheumatism) recommended adding a multiplier factor of 1.5 in the risk algorithms
6 used for patients with RA. It has been tested during the study, but it did not lead to a
7 significant improvement (**Errore. L'origine riferimento non è stata trovata.**). Overall,
8 despite an increased CV risk in case of psoriatic arthritis has been demonstrated, all the five
9 algorithms underestimated the CV risk in this study. In particular, data showed a poor model
10 fit with a significant different distribution of predicted and observed events for CUORE,
11 SCORE and RRS, but not for FRS and QRISK2. Indeed, QRISK2 showed the best
12 discriminative ability and calibration on the considered population.

13 This score was applied also to *patients affected by severe mental illness (SMIs)* in 2020,
14 which was obtained interesting result³⁰. In the last 10 years, many studies highlighted that
15 patients with SMIs show higher rates of metabolic syndrome and its connection with the
16 risk of developing CVDs³⁰. These patients are twice times more affected by diabetes,
17 obesity, and metabolic syndrome than the general population, and consequently more prone
18 to develop CV diseases. Moreover, they are normally prescribed complex medication
19 regimens and tend to be more inactive and keep unhealthy dietary habits, no matter the
20 socioeconomic status. The study, on a sample of Northern Italian patients aged 35-69 years,
21 employed the CUORE algorithm. It considered the risk in these patients compared with the
22 data from the general population, defining SMI as schizophrenia or bipolar disorder. The
23 analysis showed CV risk raise with ageing in both sexes, comparable risk in both patients
24 with low or high educational level, and it transpired those women with SMI were
25 consistently more at risk of experience a CV event than the general population counterpart,
26 even at a younger age (about 40% increased risk). Such results are supported by the higher
27 number of smoked cigarettes, the higher triglycerides, and lower HDL cholesterol levels
28 displayed in females with SMIs compared to the control group of healthy individuals. Men
29 with SMIs showed higher triglycerides, lower HDL cholesterol, and more cigarette smoking
30 than healthy subjects. However, the increased risk in males is likely offset by lower SBP in
31 patients with SMI. The CUORE have been validated on the general Italian population, but
32 it would be appropriate to incorporate specific mental health measures.

1 It emerges that the CUORE algorithm keeps an overall well fit to studies based on the Italian
 2 population, and good accuracy in presence of a bunch of comorbidities such as psoriasis or
 3 severe mental illnesses, but keeps the limitations related to a fixed function on a specific
 4 population with a limited number of variable and needs to be reviewed time by time.

5 **2 Machine Learning Approaches**

6 In the following sections, a general description and the main clinical applications of the
 7 most used ML and DL methods to predict cardiovascular risk is provided. A short
 8 description of these approaches is shown in Table VII.

9
 10

Table VII. Different ML algorithms.

<i>Logistic regression</i>	One of the most commonly used linear methods. The result is the probability that an object belongs to a specific class. Very common in medical applications. It is usually not prone to overfitting.
<i>K-nearest neighbours (KNN)</i>	Supervised learning algorithm used for objects' classification based on the characteristics of nearby objects; the input consists of k examples of training items closest in space. When the new test data is provided, KNN calculates the distance between the query data and the training samples. The Euclidean distance metric is often used. It is simple to use and implement, and it is applicable to both classification and regression problems.
<i>Probabilistic models (Naïve Bayes)</i>	A non-structural method used to demonstrate the existence of a prescribed type of mathematical object. It works by showing that if one randomly chooses objects from a specified class, the probability that the result is of the prescribed type is much greater than zero. Even if the demonstration uses probability, the conclusion is determined for certain, without any possible error.
<i>Adaptive boosting (AdaBoost)</i>	Statistical classification meta-algorithm that can be used combined with other learning algorithms; it considers as training data only samples that can improve the predictive power of the model, improving executing time avoiding the calculation of irrelevant subsets of the original dataset.
<i>Random forest</i>	An algorithm that uses a set of decision trees and the bootstrap method, which allows you to create many examples of the same size based on some randomly selected initial data and repeat; each decision tree is

	based on one of these examples. It usually presents high accuracy, but it is also often prone to overfitting.
<i>Deep learning (neural networks)</i>	The statistical model is composed of artificial neurons, inspired by a simplification of biological neural networks.

1 2.1 Logistic regression

2 Logistic regression was used for the prediction of cardiovascular events in *hypertensive*
3 *Indonesian patients*. A major health problem in Indonesia is hypertension, due to its
4 incidence at 25,8% among the population³⁴. In this study, the risk variables of Surabaya
5 Hajj hospital, such as age, heart rate, stress levels, body mass index were considered
6 influential on hypertension through a regression curve without a specific pattern. In this
7 analysis, an additive nonparametric logistic regression approach based on local linear
8 estimators was implemented to model the case of hypertension from the above-mentioned
9 variables³⁵. The non-parametric regression model was used to determine the relationship
10 between the response variable (y) and the predictive variable (t). The results of the study
11 highlighted that hypertension was most predominant among people over 65 years with BMI
12 around 25-30 kg/m² (condition of obesity) and normal heart rate around 60-100 bpm and
13 most of them were living under stressful conditions. The model obtained an accuracy of in-
14 sample data of 95% and 89,47% in the case of out-sample data. As a consequence, the
15 additive nonparametric logistic regression resulted the best estimator for determining
16 hypertension incidence. Furthermore, previous research showed a comparison between the
17 classification accuracy of hypertension using binary logistic regression and the C4.5
18 algorithm with case studies in UPT Puskesmas Ponjong I, Gunungkidul³⁴. Binary logistic
19 regression is a method that describes the relationship between the response variable and
20 multiple predictor variables with the variable equal to 1 to declare the existence of a feature
21 and the value 0 to declare the absence of a feature. In this ML method, the factors
22 influencing the onset of hypertension were sex, systolic blood pressure, and other diseases.
23 The C4.5 algorithm is a data mining classification method used to build a decision tree.
24 Factors influencing the onset of hypertension in this other technique were systolic blood
25 pressure, medical history, other diseases, diastolic blood pressure, and sex, while the age
26 factor did not affect it. Based on this analysis, the classification of hypertension by the
27 binary logistic regression method obtained an accuracy of 72.53% and the algorithm C4.5
28 obtained an accuracy of 64.08%. In conclusion, the binary logistic regression was better
29 than the C4.5 algorithm in classifying the type of hypertension.

1 The logistic regression was also used to predict the cardiovascular events in *hemodialysis*
2 *patients*, where the incidence rate is about 30 times higher than in the general population.
3 Several research groups have recently made several attempts to predict the cardiovascular
4 outcome of dialysis patients (<http://hdl.handle.net/10589/83083>). For this purpose, a dataset
5 of real data extracted from the EuCliD system, owned by the Fresenius Medical Care
6 company, was used. The data of more than 4500 patients undergoing hemodialysis treatment
7 three times a week for 18 months were analysed, in particular patients treated for the first
8 time in their life with hemodialysis treatment. This category of patients was selected for the
9 high onset of cardiovascular disease during the initial period of dialysis. The ML techniques
10 chosen for the development of these models were logistic regression based on the Lasso
11 algorithm, random forests, support vector machines, and self-organizing maps (a type of
12 side-connected neural network). The best predictive performance was obtained with the use
13 of the random forests model: a ROC AUC equal to 73% with sensitivity greater than 70%
14 was obtained, demonstrating the ability of random forests to exploit non-linear relationships
15 identified in the feature space. Two different feature selection methods both based on a
16 wrapper strategy were integrated for the construction of the models to identify subgroups of
17 features effective in prediction³⁶. From the analysis of the implemented models and the
18 analysis of the subgroups of features identified, it was possible to note that the presence of
19 an inflammatory state, malnutrition and incorrect ultrafiltration of the patients were
20 significant predictors of the onset of cardiovascular events in patients with hemodialysis
21 treatment. These factors highlighted a greater risk of a sudden deterioration in the conditions
22 of the cardiovascular system: therefore, personalized treatment strategies can be devised to
23 decrease the cardiovascular risk in patients who have suffered dialysis treatment and are
24 identified as being at greater cardiovascular risk.

25 **2.2 K-nearest neighbours (KNN)**

26 The KNN was used for the *detection and localization of MI*³⁷. The presence or absence of
27 MI is characterized by specific waves or segments in the ECG beats. The main indicators
28 are the amplitude of the T wave, the amplitude of the Q wave, and the elevation or
29 depression of the ST level. The location of the MI was predicted by analysing 12 ECG leads.
30 Using the KNN method, a subdivision into 11 classes was made for localization: 1 normal
31 class and 10 classes with different MI positions, taken from the ECG leads. The results
32 showed good sensitivity and specificity values for MI detection, 99,97% and 99,9%
33 respectively. To reduce the computational cost of the KNN classifier, a pruning algorithm

1 was employed. This method decreased the accuracy of both MI detection and localization.
2 An important aspect to consider is the control of the robustness of the feature extraction
3 because the detection and location of the MI strongly depend on the correct extraction of
4 the time-domain characteristics.

5 The KNN was employed also for the *prognosis of cardiac ischemia*, which can be detected
6 from the morphological characteristics derived from the observation of the ECG, in
7 particular from the deviation of the ST segment and the changes in the T wave³⁸. Good
8 values of sensitivity, specificity, and accuracy were achieved. In this study other approaches
9 were tried: Artificial Neural Network and Support vector machine. The experimental results
10 confirmed that the ANN model outperformed with an accuracy of 96,62%, higher compared
11 to SVM and KNN classifiers³⁸.

12 **2.3 Probabilistic methods (Naïve Bayes)**

13 Probabilistic models are used in many fields. Naïve Bayes (NB) methods are a set of
14 supervised learning algorithms based on applying Bayes' theorem with the "naive"
15 assumption of conditional independence between every pair of features given the value of
16 the class variable. This approach does not depend on the assumption that hazards are
17 proportional. For this reason, NB might be better suited to model data that do not follow a
18 proportional hazards model, or for which the functional relationship between covariates and
19 the hazard is not specified³⁹. The problem with standard NB lies in the poorly calibration
20 ability in some settings, and that is why it was necessary to extend it, developing other better
21 methods based on NB. Among the various methods of this type, one has been developed
22 called Censored Naive Bayes (CNB), which is nonparametric concerning the underlying
23 distribution of event times and models the marginal covariate distributions in a flexible way.
24 In many studies, some patients have incomplete data due to loss of follow-up or censorship
25 and ignoring or assigning a particular event status to these subjects can lead to biased
26 probability estimates. CNB is a method that considers this problem. However, CNB is
27 poorly calibrated when the covariates are not mutually independent. For this reason, it has
28 been developed another method called Censored Naive Bayes-Principal Components
29 (CNB-PC), which is a modification of CNB which results in a much better-calibrated
30 technique. In fact, CNB and CNB-PC methods were applied to *predict the risk of CVDs*
31 *from electronic health record data*⁴⁰. The data, which came from a healthcare system in the
32 Midwestern United States and were extracted from the HMO Research Network Virtual
33 Data Warehouse (HMORN VDW) associated with that system, were as follows: insurance

1 enrollment, demographics, pharmaceutical dispensing, utilization, vital signs, laboratory,
2 census, and death records. The analysis was restricted to those subjects with two medical
3 encounters in the in-network clinic with blood pressure information at least 30 days but at
4 most 1.5 years apart, with drug coverage at the end of the baseline period. Patients under
5 the age of 40 were excluded. Subjects with pre-existing serious comorbidities other than
6 diabetes were excluded too. The final analysis dataset contained 87363 individuals. The
7 cohort was randomly split into a training dataset (65522 subjects), and a test dataset (21841
8 subjects). From the results of this study, it was evident that CNB was rather poorly
9 calibrated, overpredicting the risk, whereas CNB-PC had better performances⁴⁰. In this
10 study, Cox proportional hazards (CPH) was also used to compare the performances. The
11 model emulates a situation where the effect of age on risk is nonlinear: subjects at middle
12 age are at highest risk, and those at low and high age are at lower risk. This pattern is
13 observed modelling the probability of nonfatal CV events. The additional flexibility of the
14 CNB models allowed obtaining more accurate predictions when relevant covariates were
15 omitted from the CPH model⁴¹.

16 Another method based on NB is Gaussian Naïve Bayes (GNB), which was used for a
17 classification problem and the results seemed to be more promising in terms of heart failure
18 (HF) risk prediction accuracy than other techniques in the literature. This study was
19 conducted on a sample of 297 people, and the following parameters were considered: age,
20 sex, chest pain type, resting blood pressure, serum cholesterol, fasting blood sugar, resting
21 electrocardiographic results, maximum heart rate achieved, exercise-induced angina, old
22 peak, peak exercise slope, number of major vessels coloured by fluoroscopy and thallium
23 scan. The robustness of the proposed model was evaluated using accuracy, specificity, and
24 sensitivity as evaluation metrics and a forward feature selection. The best model was
25 achieved with a subset of 9 features (including sex, chest pain type, resting
26 electrocardiographic results, maximum heart rate achieved, exercise-induced angina, old
27 peak, peak exercise slope, number of major vessels coloured by fluoroscopy, thallium scan)
28 with an accuracy on test of 93%, a training accuracy of 84%, a sensitivity of 88% and a
29 specificity of 98%. The same study was conducted using the RF method, to make a
30 comparison between RF and GNB and the latter performed better. It is possible to verify the
31 greater effectiveness of GNB compared to RF also by observing the respective ROC AUC
32 of 96% and 94%, respectively. The GNB approach outperformed the RF approach⁴².

1 **2.4 Tree-based models**

2 **2.4.1 Adaptive boosting**

3 In 2021 D'Ascenzo et al. conducted a study to compare the performance of 4 ML methods
4 and identify the best model to predict the *risk of death, MI, and risk of bleeding 1 year after*
5 *discharge*⁴³. A sample of 23270 adult patients (derivation cohort) and an additional sample
6 of 3444 patients (external validation cohort) were analysed. The derivation cohort was
7 randomly split into a training dataset (80%) and a validation dataset (20%). The analysed
8 methods were AB, NB, RF, and KNN. AB, as known as the PRAISE model, was found to
9 be the most performing model for each of the three cases. The PRAISE model considers 25
10 variables (clinical, anatomical, and procedural characteristics). A subset of the 25 features
11 was found to be important for the three types of risk prediction separately. The most
12 significant variables for the prediction of death are LVEF, age, hemoglobin level, statin (at
13 discharge), EGFR, ACE inhibitor or ARB, previous bleeding, and malignancy. The main
14 variables for the prediction of acute MI are hemoglobin level, LVEF, age, EGFR, peripheral
15 artery disease, diabetes, complete revascularisation, multivessel disease. The main features
16 for the prediction of risk of bleeding are hemoglobin level, age, LVEF, EGFR, diabetes,
17 drug-eluting stent, complete revascularisation, malignancy. The PRAISE model classified
18 60% of acute coronary syndrome (ACS) patients as at low risk of ischemic and hemorrhagic
19 events (1% probability) and 10% of ACS patients as at high risk of these events (>19%).
20 The above-mentioned percentages are similar to those observed in the clinical setting.
21 Among the considered models, the PRAISE model is the one with the most performing
22 ROC curve in all of the three cases and accuracy, sensitivity, specificity, and F2 score values
23 of 89%, 55%, 91%, and 43% on the evaluation set and 78%, 88%, 77% and 60% on the
24 validation set, respectively⁴³. The limitations of this approach are the slight underestimation
25 of the adaptive enhancement classifier among high-risk patients and the no inclusion of the
26 time as a variable since it has been created to be used with a fixed time frame (e.g. 1-year
27 follow-up)⁴³.

28 **2.4.2 Random forest**

29 RF methods represent one of the most important families of ML methods based on the use
30 of decision trees. This ML method is used in the CVD prediction setting due to its enhanced
31 predictive ability and the ability to find linear and non-linear interactions between
32 parameters and risk factors. This could improve the prediction in presence of previous

1 diseases. Finally, RF can be used in all applications involving a large group of data, in fact,
2 the reliability of the algorithm is not compromised by the great number of data^{44,45}. ML's
3 unique feature is that algorithms do not require any assumptions on the underlying data;
4 therefore, non-linearities can be caught more easily. Since considered risk factors often
5 result insufficient to analyse complex relationships between heterogeneous data, this
6 explains the necessity to implement ML methods in risk scores to underline the weight of
7 specific illnesses on cardiovascular risk. Some of these methods were implemented in the
8 study of inflammatory diseases using C-reactive protein as a marker of an inflammatory
9 state. The use of the RF method was validated in the prediction of CVD in patients affected
10 by Psoriatic Arthritis (PsA) and the results showed a higher sensitivity in presence of this
11 kind of disease⁴⁶. The results obtained with PsA generated an AUC of 83% which is
12 significantly higher than the value of AUC obtained with the use of different ML methods
13 as the KNN. The use of the RF was studied also in presence of other diseases as Ankylosing
14 Spondylitis (AS), an inflammatory disease that can be localized in vertebral articulations.
15 In this case, the RF method was implemented to predict the cardiovascular risk in AS
16 patients and was compared with other ML methods and traditional risk scores⁴⁷. The results
17 showed better accuracy and specificity, of 65% and 66% respectively, and a lower
18 sensitivity of 61% of the RF method if compared to the KNN method. However, the best
19 algorithms that permit the prediction of CV risk resulted to be Reynolds Risk Score (RRS)
20 and SCORE which have a higher specificity, accuracy, and sensitivity than ML methods.
21 Finally, the most important limitations of this study are: first, the fact that the cohort of
22 people suffering from AS involves a small number of subjects and therefore a large
23 validation set for ML methods like RF, that need a significant number of data to be correctly
24 balanced, is not available; second, the entire cohort sample is composed only by Italian
25 Caucasian people.

26 **2.5 Deep learning (neural networks)**

27 Artificial neural networks (ANN) were used in *HF prediction using cardiopulmonary*
28 *exercise testing (CPX)*, which was employed to quantify the degree of exercise intolerance
29 patients with chronic HF. Recently many studies demonstrated how CPX can be useful in
30 determining the risk of mortality, hospitalization, and other adverse outcomes in subject
31 with HF⁴⁸, and how peak VO₂, ventilatory abnormalities, heart rate recovery (HRR), and
32 other responses, could provide clinically significant and independent information for
33 estimating prognosis in these patients. ANNs overcome the limitation of the traditional

1 statistical methods, providing a more accurate value for the risk estimation⁴⁸. The study
2 aimed to evaluate the predictive accuracy of the chosen CPX using an ANN and compare it
3 to conventional methods to estimate the risk of CV mortality. The dataset included 2635
4 patients, which were split into training and testing subsets, therefore the 2635 recruited
5 patients, 550 were not included in the training set since they have been used for testing the
6 ANN. The subjects were followed for a major cardiac event for three years. ANN analyses
7 were performed using a feed-forward multilayer learning strategy with a back-propagation
8 training method. ROC curves were evaluated to compare the predicting value of each
9 method. In the follow-up, it emerged that 291 subjects died from cardiovascular causes, and
10 their data have been compared with the ones from the survivors. Each component of the
11 CPX resulted significantly connect to CV mortality. Through logistic regression and
12 proportional hazard, it was possible to determine the predictive strength of the observed
13 variables (VE/VCO₂ slope revealed to be the strongest predictor of risk), while ANN
14 showed higher predictive accuracy using 3 clinical variables together with the 5 CPX
15 variables. ANN exhibited a better performance compared to the other two more traditional
16 methods with an AUC of 72%, a sensitivity of 79% and a specificity of 63%. Moreover,
17 ANN reached its ROC area peak at 5 hidden neurons, after which it decreased due to
18 overfitting. However, no significant differences in the ROC areas in the three methods were
19 observed. This study was applied to a cohort of 75% male, so it may not represent a female
20 population, and the performance was evaluated using few variables, which could be the limit
21 of this study.

22 Myers and co-workers applied the sparse autoencoder (SAE) approach to perform
23 unsupervised learning for heart disease prediction⁴⁹. SAE is an unsupervised neural
24 network, which gives in output a model reconstructed through backpropagation of the input
25 data and the reconstruction error, difference of the original input and its preconstruction.
26 The dataset consisted of 4238 samples and 16 features, refers to a cardiovascular study on
27 residents of Framingham, in Massachusetts, predicting their 10-years risk of CHD. Due to
28 missing data, only 3656 have been considered, 3099 negative and 557 positive. The authors
29 did a comparison between the ANN and SAE and they noticed that the combination of the
30 two methods performed better. Then, comparative experiments were carried out using some
31 of the algorithms proposed in this review to verify the effectiveness of the model. In
32 particular, there was a comparison between the above-mentioned method and the
33 performance of KNN, Classification and Regression Trees (CART), Logistic Regression

1 (LR), Naïve Bayes. The above-mentioned method reported the best performances with 90%
2 of accuracy, 89 % of precision, 91% of recall and 90% of F1-score⁴⁹.

3 Traditionally, congestive heart failure (CHF) detection with ML on ECG relied on the
4 extraction of heart rate variability features to build classifiers able to discriminate between
5 healthy and CHF subjects. Porumb et al. proposed a new framework of congestive heart
6 failure detection relying on raw ECG signals, through a 1-D Convolutional Neural Network,
7 which is another type of neural network in DL⁵⁰. Working on raw physiological data
8 improves the interpretability of the model and ensures the transparency of the method⁵⁰. A
9 retrospective analysis was performed on two public datasets: the MIT-BIH Normal Sinus
10 Rhythm Database from PhysioNet for the control group, including 18 healthy non-
11 arrhythmic subjects for a total of 275974 beats considered, and the BIDMC Congestive
12 Heart Failure Database from PhysioNet for the CHF group including 15 pathological
13 subjects, for a total of 214531 beats considered. About 20 hours of recording were available
14 for each subject, consequently, due to the very large amount of data collected, it was
15 randomly selected one single beat every 5 seconds. All data are from the same laboratory.
16 Total data was randomly split in smaller subsets as follows: 50% for training, 25% for
17 validation and 25% for testing. Random splitting was repeated 10 times to reduce variance.
18 Results showed the number of false positive is about 1% of the mean number of normal
19 heartbeats, and the false negative are around 3% of the total number of heartbeats in CHF
20 class. 478 from the control group and 854 CHF beats were misclassified, in both cases
21 mainly belonging to the same subject. Since more than the 95% of the heartbeats were
22 properly classified, patients could still be correctly categorized. Results have been
23 confirmed by an additional simulation averaging the heartbeats in 5 s interval. No changes
24 was detected in the performance. Employing the majority voting scheme for heartbeats
25 within 5 minutes of ECG, it was possible to evaluate the accuracy of the method in the
26 discrimination of the patients from the control or CHF group. Results suggest that it was
27 possible to build a highly accurate model for CHF diagnosis using an about 5 minutes long
28 raw ECG recording. Authors showed, indeed, that 99% of the 5 minutes ECG extracts could
29 be employed for providing a correct CHF diagnosis. The limitation of this study is the
30 limited dataset and it is needed to extend the obtained results to larger datasets.

31 López-Martínez et al. presented an alternative approach using ANN to classify hypertensive
32 patients⁵¹. The dataset contains patients from the National Health and Nutrition Examination
33 Survey (NHANES) from 2007 to 2016. Data are public and collected under survey form,

1 combining also social determinants of health data such as smoking, alcohol, and diet. This
2 dataset is very unbalanced since the survey involved a generic population, not a specific
3 one. Despite an acceptable model accuracy of 73%, the disbalance of the dataset affects the
4 precision and sensitivity of the model which reveals a high number of false negatives. Some
5 other ML algorithms were employed with the same NHANES data to compare their AUC
6 and to validate the performance and sensitivity of this method in disease diagnostic. All the
7 methods show very similar accuracy working on the same inaccurate dataset, whereas AUC
8 and F1-score result was slightly higher in the proposed model. ANN with back-propagation
9 was used to classify hypertensive patients. Out of 7330 patients included in the dataset,
10 69,9% were non-hypertensive and 30,1% were hypertensive. Therefore, the model shows a
11 sensitivity of 40% and a specificity of 87%. The predictive power of positives and negatives
12 was respectively 57% and 77%, with 12% of false alarm cases. Such results suggest that the
13 proposed ANN model might not fit positive detection in healthcare diagnosis but are
14 conversely effective in finding non-hypertensive patients with an accuracy of 73%. In
15 conclusion, this model cannot be used for giving a final diagnosis but can be used to make
16 clinicians aware of the probability that the patients could develop hypertension⁵¹.

17 **CONCLUSION**

18 In the present review, two different approaches to the prevention of cardiovascular risk were
19 described. According to the considered risk factors and the specific cardiovascular event, it
20 could be convenient to move towards either traditional methods or more innovative methods
21 of ML, as one predictive model can be more effective than another.

22 The main difference between ML and traditional methods is that ML learns from data and
23 is dependent on training data, considering a greater number of both linear and non-linear
24 variables, while statistical models use a limited number of variables only with a linear
25 relationship between events and risk factors. ML algorithms are often highly flexible
26 algorithms that require a penalty to avoid overfitting. Therefore, ML algorithms can take
27 into account much more information⁵². In studies where the main objective is to predict the
28 event without considering the association between specific risk factors and a clinically
29 understandable event, traditional methods are replaced by the use of ML, since traditional
30 methods are often relevant and necessary in association analyses but do not fit in prediction
31 analyses, where the focus is on the output and not on the initial variables. ML and traditional
32 methods often have the same objective, but are built differently: ML does not need pre-
33 specified basic structures, but rather looks for the best fit: at the end, a better predictive

1 model is generated, at the expense of understanding how the risk factors are related to the
2 event of interest. In particular, the final goal is always to perform accurate and precise
3 predictions, being robust to future applications (e.g., validation on external datasets) and to
4 elucidate the unknown relationships between data. Traditional statistical methods are based
5 on theoretical and a priori hypotheses and assumptions and they draw inference from
6 predictors. ML methodologies are instead data-driven models empirically optimized based
7 on agnostic approaches which enable exploratory data analysis and insights. In addition,
8 traditional methods are based on regression models that suppose a linear relationship
9 between the event and risk factors, while ML can consider the non-linearity of some risk
10 factors giving a more precise result (e.g., the risk of death increases with age, but the
11 increase seen in the 40-50 age group is different from that in the 70-80 age group). At the
12 same time, however, an algorithm could have reduced performance if applied, for example,
13 to patients belonging to ethnic minorities if these subjects were not included inside the
14 training phase of the algorithm.

15 Thanks to its flexibility, it is believed that ML in the future will have better performance
16 than traditional statistical modelling. The prediction of cardiovascular risk is increasingly
17 approaching an automated sphere controlled by AI. Despite all the difficulties of using ML
18 approaches, this will greatly facilitate the work of doctors, increase the effectiveness of
19 diagnostics and contribute to the choice of effective treatment measures.

20 **KEY MESSAGES**

- 21 • *Traditional Risk Score methods* are based on theoretical and a priori hypotheses and
22 assumptions. They can catch the linear relationship between the event and risk
23 factors, as they are regressor models.
- 24 • *Machine Learning methods* are data-driven models based on agnostic approaches
25 that allow for explorative data analysis. They can catch the non-linearity between
26 risk factors and events giving a more accurate risk prediction. The training of the
27 model can be affected by the data.
- 28 • Machine Learning methods can improve the CVD risk prediction supporting clinical
29 decision and early diagnosis.

30

31

REFERENCES

- 32 1. Abidov A, Chehab O. Cardiovascular risk assessment models: Have we found the

- 1 perfect solution yet? *J Nucl Cardiol* [Internet]. 2020 Dec 21;27(6):2375–85.
2 Available from: <https://doi.org/10.1007/s12350-019-01642-x>
- 3 2. Ming C, Viassolo V, Probst-Hensch N, Chappuis PO, Dinov ID, Katapodi MC.
4 Machine learning techniques for personalized breast cancer risk prediction:
5 comparison with the BCRAT and BOADICEA models. *Breast Cancer Res*
6 [Internet]. 2019 Dec 20;21(1):75. Available from: [https://breast-cancer-](https://breast-cancer-research.biomedcentral.com/articles/10.1186/s13058-019-1158-4)
7 [research.biomedcentral.com/articles/10.1186/s13058-019-1158-4](https://breast-cancer-research.biomedcentral.com/articles/10.1186/s13058-019-1158-4)
- 8 3. Churpek MM, Carey KA, Edelson DP, Singh T, Astor BC, Gilbert ER, et al.
9 Internal and External Validation of a Machine Learning Risk Score for Acute
10 Kidney Injury. *JAMA Netw Open* [Internet]. 2020 Aug 11;3(8):e2012892.
11 Available from:
12 <https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2769146>
- 13 4. Halasz G, Sperti M, Villani M, Michelucci U, Agostoni P, Biagi A, et al. A
14 Machine Learning Approach for Mortality Prediction in COVID-19 Pneumonia:
15 Development and Evaluation of the Piacenza Score. *J Med Internet Res* [Internet].
16 2021 May 31;23(5):e29058. Available from: <https://www.jmir.org/2021/5/e29058>
- 17 5. Firyulina MA, Kashirina IL. Using of machine learning methods in cardiology. *J*
18 *Phys Conf Ser* [Internet]. 2021 May 1;1902(1):012135. Available from:
19 <https://iopscience.iop.org/article/10.1088/1742-6596/1902/1/012135>
- 20 6. Colpani O, Preventiva F, Farmacologiche S. Machine learning : the ability to
21 predict applied to research and clinical practice. 2019;11(4):5–11.
- 22 7. Andersson C, Johnson AD, Benjamin EJ, Levy D, Vasani RS. 70-year legacy of the
23 Framingham Heart Study. *Nat Rev Cardiol* [Internet]. 2019;16(11):687–98.
24 Available from: <http://dx.doi.org/10.1038/s41569-019-0202-5>
- 25 8. D’Agostino RB, Vasani RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, et al.
26 General cardiovascular risk profile for use in primary care: The Framingham heart
27 study. *Circulation*. 2008;117(6):743–53.
- 28 9. Orfanoudaki A, Chesley E, Cadisich C, Stein B, Nouh A, Alberts MJ, et al. Machine
29 learning provides evidence that stroke risk is not linear: The non-linear
30 Framingham stroke risk score. *PLoS One* [Internet]. 2020;15(5):1–20. Available
31 from: <http://dx.doi.org/10.1371/journal.pone.0232414>
- 32 10. Benjamin EJ, Wolf PA, D’Agostino RB, Silbershatz H, Kannel WB, Levy D.
33 Impact of atrial fibrillation on the risk of death: The Framingham Heart Study.

- 1 Circulation. 1998;98(10):946–52.
- 2 11. Batty GD, Hamer M. Vascular risk factors, framinghamrisk score, and covid-19:
3 Community-based cohort study. *Cardiovasc Res.* 2020;116(10):1664–5.
- 4 12. DAWBER TR, MEADORS GF, MOORE FE. Epidemiological approaches to heart
5 disease: the Framingham Study. *Am J Public Health.* 1951;41(3):279–81.
- 6 13. Cuende JI, Cuende N, Calaveras-Lagartos J. How to calculate vascular age with the
7 SCORE project scales: A new method of cardiovascular risk evaluation. *Eur Heart*
8 *J.* 2010;31(19):2351–8.
- 9 14. Goh LGH, Welborn TA, Dhaliwal SS. Independent external validation of
10 cardiovascular disease mortality in women utilising Framingham and SCORE risk
11 models: A mortality follow-up study. *BMC Womens Health.* 2014;14(1):1–11.
- 12 15. Piccininni M, Rohmann JL, Huscher D, Mielke N, Ebert N, Logroscino G, et al.
13 Performance of risk prediction scores for cardiovascular mortality in older persons:
14 External validation of the SCORE OP and appraisal. *PLoS One.* 2020;15(4):1–18.
- 15 16. Fernández-Labandera C, Calvo-Bonacho E, Valdivielso P, Quevedo-Aguado L,
16 Martínez-Munoz P, Catalina-Romero C, et al. Prediction of fatal and non-fatal
17 cardiovascular events in young and middle-aged healthy workers: The IberScore
18 model. *Eur J Prev Cardiol.* 2019;
- 19 17. Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, Brindle P. Performance
20 of the QRISK cardiovascular risk prediction algorithm in an independent UK
21 sample of patients from general practice: A validation study. *Heart.* 2008;94(1):34–
22 9.
- 23 18. Collins GS, Altman DG. An independent and external validation of QRISK2
24 cardiovascular disease risk score: A prospective open cohort study. *BMJ.*
25 2010;340(7758):1231.
- 26 19. Collins GS, Altman DG. Predicting the 10 year risk of cardiovascular disease in the
27 United Kingdom: Independent and external validation of an updated version of
28 QRISK2. *BMJ.* 2012;345(7867):1–12.
- 29 20. Finnikin S, Ryan R, Marshall T. Statin initiations and QRISK2 scoring in general
30 practice. 2017;(December). Available from:
31 <https://doi.org/10.3399/bjgp17X693485>
- 32 21. Hippisley-Cox J, Coupland C, Robson J, Brindle P. Derivation, validation, and
33 evaluation of a new QRISK model to estimate lifetime risk of cardiovascular

- 1 disease: Cohort study using QResearch database. *Bmj*. 2011;342(7788):93.
- 2 22. Assmann G, Cullen P, Schulte H. Simple scoring scheme for calculating the risk of
3 acute coronary events based on the 10-year follow-up of the Prospective
4 Cardiovascular Münster (PROCAM) study. *Circulation*. 2002;105(3):310–5.
- 5 23. Cox DR. Regression Models and Life-Tables. *J R Stat Soc Ser B*. 1972;34(2):187–
6 202.
- 7 24. Batsis JA, Sarr MG, Collazo-Clavell ML, Thomas RJ, Romero-Corral A, Somers
8 VK, et al. Cardiovascular Risk After Bariatric Surgery for Obesity. *Am J Cardiol*.
9 2008;102(7):930–7.
- 10 25. Adams TD, Gress RE, Smith SC, Halverson RC, Simper SC, Rosamond WD, et al.
11 Long-Term Mortality After Gastric Bypass Surgery. *Surv Anesthesiol* [Internet].
12 2008 Apr;52(2):94. Available from: [http://journals.lww.com/00132586-200804000-](http://journals.lww.com/00132586-200804000-00040)
13 00040
- 14 26. Cook NR, Paynter NP, Eaton CB, Manson JE, Martin LW, Robinson JG, et al.
15 Comparison of the framingham and reynolds risk scores for global cardiovascular
16 risk prediction in the multiethnic women’s health initiative. *Circulation*.
17 2012;125(14):1748–56.
- 18 27. Buring JE, Cook NR. Algorithms for the Assessment of Global Cardiovascular Risk
19 in Women - The Reynolds Risk Score. *Am Med Assoc*. 2016;297(6):611–20.
- 20 28. Ridker PM, Silvertown JD. Inflammation, C-Reactive Protein, and
21 Atherothrombosis. *J Periodontol*. 2008;79(8s):1544–51.
- 22 29. Doukaki S, Caputo V, Bongiorno MR. Psoriasis and cardiovascular risk:
23 Assessment by cuore project risk score in Italian patients. *Dermatol Res Pract*.
24 2013;2013.
- 25 30. Salvi V, Aguglia A, Barone-Adesi F, Bianchi D, Donfrancesco C, Dragogna F, et
26 al. Cardiovascular risk in patients with severe mental illness in Italy. *Eur*
27 *Psychiatry*. 2020;63(1):4–9.
- 28 31. Zanchè A, Lapi F, Parretti D. Cardiovascular risk chart : past , present and future.
29 *Rivista SIMG*. 2021;28(1):32–7.
- 30 32. La Torre G, Palmeri V, Pagano L, Nati G, Sernia S, Mannocci A. Evaluation of
31 cardiovascular risk profile: A comparative analysis between CUORE algorithm and
32 the Framingham risk scores. *Eur Rev Med Pharmacol Sci*. 2018;22(8):2398–404.
- 33 33. Navarini L, Margiotta DPE, Caso F, Currado D, Tasso M, Angeletti S, et al.

- 1 Performances of five risk algorithms in predicting cardiovascular events in patients
2 with Psoriatic Arthritis: An Italian bicentric study. *PLoS One*. 2018;13(10):1–11.
- 3 34. Rumaenda W, Wilandari Y, Safitri D. Perbandingan Klasifikasi Penyakit Hipertensi
4 Menggunakan Regresi Logistik Biner Dan Algoritma C4.5 (Studi Kasus Upt
5 Puskesmas Ponjong I, Gunungkidul). *J GAUSSIAN* [Internet]. 2016;5(2):299–309.
6 Available from: <http://ejournal-s1.undip.ac.id/index.php/gaussian>
- 7 35. Ana E, Chamidah N, Andriani P, Lestari B. Modeling of hypertension risk factors
8 using local linear of additive nonparametric logistic regression. *J Phys Conf Ser*.
9 2019;1397(1).
- 10 36. Dimensioni P, Di C, Di T. FEATURE SELECTION PER Introduzione e contesto.
11 2017;(April).
- 12 37. Arif M, Malagore IA, Afsar FA. Detection and localization of myocardial infarction
13 using K-nearest neighbor classifier. *J Med Syst*. 2012;36(1):279–89.
- 14 38. Murthy HSN, Meenakshi DM. ANN, SVM and KNN Classifiers for Prognosis of
15 Cardiac Ischemia- A Comparison. *Bonfring Int J Res Commun Eng*. 2015;5(2):07–
16 11.
- 17 39. Webb GI, Keogh E, Miikkulainen R, Miikkulainen R, Sebag M. Naïve Bayes. In:
18 *Encyclopedia of Machine Learning* [Internet]. Boston, MA: Springer US; 2011. p.
19 713–4. Available from: http://link.springer.com/10.1007/978-0-387-30164-8_576
- 20 40. Wolfson J, Bandyopadhyay S, Elidrisi M, Vazquez-Benitez G, Vock DM,
21 Musgrove D, et al. A Naive Bayes machine learning approach to risk prediction
22 using censored, time-to-event data. *Stat Med* [Internet]. 2015 Sep 20;34(21):2941–
23 57. Available from: <http://doi.wiley.com/10.1002/sim.6526>
- 24 41. Wolfson J, Bandyopadhyay S, Elidrisi M, Vazquez-Benitez G, Vock DM,
25 Musgrove D, et al. A Naive Bayes machine learning approach to risk prediction
26 using censored, time-to-event data. *Stat Med*. 2015;34(21):2941–57.
- 27 42. Ali L, Khan SU, Golilarz NA, Yakubu I, Qasim I, Noor A, et al. A Feature-Driven
28 Decision Support System for Heart Failure Prediction Based on χ^2 Statistical Model
29 and Gaussian Naive Bayes. *Comput Math Methods Med*. 2019;2019.
- 30 43. D’Ascenzo F, De Filippo O, Gallone G, Mittone G, Deriu MA, Iannaccone M, et al.
31 Machine learning-based prediction of adverse events following an acute coronary
32 syndrome (PRAISE): a modelling study of pooled datasets. *Lancet*.
33 2021;397(10270):199–207.

- 1 44. Svetnik V, Liaw A, Tong C, Christopher Culberson J, Sheridan RP, Feuston BP.
2 Random Forest: A Classification and Regression Tool for Compound Classification
3 and QSAR Modeling. *J Chem Inf Comput Sci*. 2003;43(6):1947–58.
- 4 45. Yang L, Wu H, Jin X, Zheng P, Hu S, Xu X, et al. Study of cardiovascular disease
5 prediction model based on random forest in eastern China. *Sci Rep*. 2020;10(1):1–
6 8.
- 7 46. Navarini L, Sperti M, Currado D, Costa L, Deriu MA, Margiotta DPE, et al. A
8 machine-learning approach to cardiovascular risk prediction in psoriatic arthritis.
9 *Rheumatol (United Kingdom)*. 2020;59(7):1767–9.
- 10 47. Navarini L, Caso F, Costa L, Currado D, Stola L, Perrotta F, et al. Cardiovascular
11 Risk Prediction in Ankylosing Spondylitis: From Traditional Scores to Machine
12 Learning Assessment. *Rheumatol Ther [Internet]*. 2020;7(4):867–82. Available
13 from: <https://doi.org/10.1007/s40744-020-00233-4>
- 14 48. Myers J, De Souza CR, Borghi-Silva A, Guazzi M, Chase P, Bensimhon D, et al. A
15 neural network approach to predicting outcomes in heart failure using
16 cardiopulmonary exercise testing. *Int J Cardiol [Internet]*. 2014;171(2):265–9.
17 Available from: <http://dx.doi.org/10.1016/j.ijcard.2013.12.031>
- 18 49. Mienye ID, Sun Y, Wang Z. Improved sparse autoencoder based artificial neural
19 network approach for prediction of heart disease. *Informatics Med Unlocked*
20 [Internet]. 2020;18:100307. Available from:
21 <https://doi.org/10.1016/j.imu.2020.100307>
- 22 50. Porumb M, Iadanza E, Massaro S, Pecchia L. A convolutional neural network
23 approach to detect congestive heart failure. *Biomed Signal Process Control*
24 [Internet]. 2020;55:101597. Available from:
25 <https://doi.org/10.1016/j.bspc.2019.101597>
- 26 51. López-Martínez F, Núñez-Valdez ER, Crespo RG, García-Díaz V. An artificial
27 neural network approach for predicting hypertension using NHANES data. *Sci Rep*
28 [Internet]. 2020;10(1):1–14. Available from: [https://doi.org/10.1038/s41598-020-](https://doi.org/10.1038/s41598-020-67640-z)
29 [67640-z](https://doi.org/10.1038/s41598-020-67640-z)
- 30 52. Goldstein BA, Navar AM, Carter RE. Moving beyond regression techniques in
31 cardiovascular risk prediction: Applying machine learning to address analytic
32 challenges. *Eur Heart J*. 2017;38(23):1805–14.

1 *Conflicts of interest.* — The authors certify that there is no conflict of interest with any
2 financial organization regarding the material discussed in the manuscript.

3

4 *Authors' contributions.* — MS, MM and MAD have given substantial contributions to the
5 conception of the manuscript, FSP, AD, SB, AF, RR and CS to acquisition, analysis and
6 interpretation of the data. All authors have participated to drafting the manuscript, MAD
7 revised it critically. All authors read and approved the final version of the manuscript.

8 **TABLES**

9 **Table I.** Regression Coefficients and Hazard Ratios¹². SBP = systolic blood pressure.

10 **Table II.** Variables used in the construction of the Cox proportional hazards model. The P-
11 value is <0,001 for all variables.

12 **Table III.** Coefficients of PROCAM Scoring Scheme.

13 **Table IV.** 10 years risk of acute coronary events associated with each PROCAM score²².

14 **Table V.** This table shows the coefficient β and standards errors (SE) for RRS. The P-value
15 is <0,001 for all variables. Abbreviations: HbA1c = hemoglobin A1c; SBP = Systolic blood
16 pressure; HDL-C = high-density lipoprotein cholesterol; hsCRP = high- sensitivity C-
17 reactive protein; MI = myocardial infarction²⁷.

18 **Table VI.** Cox proportional hazard model for the CV assessment in 10 years³¹.

19 **Table VII.** Different ML algorithms.

20 **TITLES OF FIGURES**

21 **Figure 1.** The two main categories of risk factors: non-modifiable risk and modifiable risk.
22 Modifiable risk factors also include lifestyle factors (diet, physical exercise, stress, and
23 smoking) and partly social factors (social deprivation and environment).

24 **Figure 2.** A summary of the main pros and cons of traditional risk scores vs machine
25 learning-based approaches.