

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

Development of a Prediction Score to Avoid Confirmatory Testing in Patients With Suspected Primary Aldosteronism

This is the author's manuscript

Original Citation:

Availability:

This version is available <http://hdl.handle.net/2318/1796285> since 2021-08-09T18:58:49Z

Published version:

DOI:10.1210/clinem/dgaa974

Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

1 **Development of a prediction score to avoid confirmatory testing in patients with suspected**
2 **primary aldosteronism.**

3 Jacopo Burrello MD^{1*}, Martina Amongero MS^{2*}, Fabrizio Buffolo MD¹, Elisa Sconfienza MD¹,
4 Vittorio Forestiero MD¹, Alessio Burrello MS³, Christian Adolf MD⁴, Laura Handgriff MD⁴, Martin
5 Reincke MD⁴, Franco Veglio MD¹, Tracy Ann Williams PhD^{1,4}, Silvia Monticone MD, PhD^{1#}, Paolo
6 Mulatero MD^{1#}.

7

8 (1) Division of Internal Medicine and Hypertension, Department of Medical Sciences, University of
9 Torino, Italy. (2) Department of Mathematical Sciences G. L. Lagrange, Polytechnic University of
10 Torino, Italy. (3) Department of Electrical, Electronic and Information Engineering "Guglielmo
11 Marconi" (DEI), University of Bologna, Italy. (4) Medizinische Klinik und Poliklinik IV, Klinikum
12 der Universität, Ludwig-Maximilians-Universität München, Munich, Germany.

13 * Contributed equally and should be considered as joint first authors.

14 # Contributed equally and should be considered as joint last authors.

15

16 **Short title:** Diagnostic modelling to confirm PA diagnosis

17 **Keywords:** aldosterone; primary aldosteronism; confirmatory testing; machine learning.

18

19 **Corresponding author and person to whom reprints should be addressed:** Prof. Paolo Mulatero
20 Division of Internal Medicine and Hypertension Unit, Department of Medical Sciences, University
21 of Torino, Città della Salute e della Scienza, Via Genova 3, 10126 Torino, Italy. Telephone/Fax
22 number: +39.011.633.6959 / +39.011.633.6931. E-mail: paolo.mulatero@unito.it

23

24 **Source(s) of Funding:** This research did not receive any specific grant in Torino. The German Conn-
25 Registry-Else-Kröner Hyperaldosteronism Registry is supported by the Else Kröner-Fresenius
26 Stiftung (2013_A182, 2015_A171 and 2019_A104). CA and MR are supported by the Deutsche

27 Forschungsgemeinschaft (DFG, within the CRC/Transregio 205/1 “The Adrenal: Central Relay in
28 Health and Disease”).

29 **Conflict(s) of Interest/Disclosure(s):** the authors have nothing to disclose.

30

31 **ABSTRACT**

32 **Context.** The diagnostic work-up of primary aldosteronism (PA) includes screening and confirmation
33 steps. Case confirmation is time-consuming, expensive, and there is no consensus on tests and
34 thresholds to be used. Diagnostic algorithms to avoid confirmatory testing may be useful for the
35 management of patients with PA.

36 **Objective.** Development and validation of diagnostic models to confirm or exclude PA diagnosis in
37 patients with a positive screening test.

38 **Design, Patients and Setting.** We evaluated 1,024 patients who underwent confirmatory testing for
39 PA. The diagnostic models were developed in a training cohort (n=522), and then tested on an internal
40 validation cohort (n=174) and on an independent external prospective cohort (n=328).

41 **Main outcome measure.** Different diagnostic models and a 16-point score were developed by
42 machine learning and regression analysis to discriminate patients with a confirmed diagnosis of PA.

43 **Results.** Male sex, antihypertensive medication, plasma renin activity, aldosterone, potassium levels
44 and presence of organ damage were associated with a confirmed diagnosis of PA. Machine learning
45 based models displayed an accuracy of 72.9-83.9%. The Primary Aldosteronism Confirmatory
46 Testing (PACT) score correctly classified 84.1% at training and 83.9% or 81.1% at internal and
47 external validation, respectively. A flow chart employing the PACT score to select patients for
48 confirmatory testing, correctly managed all patients, and resulted in a 22.8% reduction in the number
49 of confirmatory tests.

50 **Conclusions.** The integration of diagnostic modelling algorithms in clinical practice may improve
51 the management of patients with PA by circumventing unnecessary confirmatory testing.

52 **INTRODUCTION**

53 Primary aldosteronism (PA) represents the most frequent cause of secondary hypertension, with a
54 prevalence reaching 29.8% in referral centers [1,2]. According to the Endocrine Society (ES), high
55 risk groups accounting for up to 50% of patients with hypertension should be screened for PA by
56 measurement of the aldosterone-to-renin ratio (ARR) [1,2]. Guidelines do not recommend an ARR
57 cut-off for a positive screening test. Nonetheless, an ARR value ranging between 20 and 40 (as
58 [ng/dL]/[ng/mL/h]) is suggested, depending on assay used for aldosterone and renin measurements
59 and inter-center variability [1,2]. The cut-off used by the majority of referral centers is 30
60 [ng/dL]/[ng/mL/h], which maximizes sensitivity, but may lead to many false positive results [3]. For
61 this reason, confirmatory testing, to either confirm or exclude the diagnosis of PA, is recommended
62 in patients with a positive case detection. In particular, the ES guideline recommends one of the
63 following: saline infusion test, captopril challenge test, oral sodium loading test, or fludrocortisone
64 suppression test [1,2]. The confirmatory test could be avoided for patients who display spontaneous
65 hypokalemia, suppressed renin levels and an aldosterone at screening greater than 20 ng/dL [1,2].
66 The Japanese Endocrine Society recommends the performance of at least 2 different confirmatory
67 tests for all patients with a positive screening test [4].

68

69 Confirmatory testing aims to identify false positives at screening to avoid subsequent costly and
70 invasive investigations, including adrenal vein sampling. However, these tests are time-consuming
71 and there is no consensus on the best test or on the thresholds that should be used for confirmation or
72 exclusion of PA. Studies that assessed between-test comparability suffer from several limitations,
73 including sample size, retrospective design, and selection bias [1,2].

74

75 Diagnostic algorithms that employ clinical and biochemical parameters at screening would be useful
76 to identify patients who can bypass confirmatory testing and proceed directly to subtype
77 differentiation due to a high likelihood of PA and those with such a low likelihood that confirmatory

78 testing is unnecessary. The aim of the present study was to develop and validate computational
79 models to confirm or exclude the diagnosis of PA in patients with a positive screening test. We
80 propose different diagnostic algorithms based on machine learning techniques, and a flow chart for
81 patient management which integrates the Primary Aldosteronism Confirmatory Testing (PACT)
82 score to stratify patients according to their likelihood of PA.

83

84 **METHODS**

85 Single patient data extracted during the present study are not publicly available but are available from
86 the corresponding author on reasonable request. Supplemental Data [5] are available at the link:
87 <https://github.com/CentroIpertenUnito/PACT-score>.

88

89 *Data extraction and study cohorts*

90 For the developmental cohort, we retrospectively assessed data from 696 patients referred to the
91 tertiary hypertension unit of Torino in whom confirmatory testing had been performed for a suspected
92 diagnosis of PA. Inclusion criteria were: (1) a positive screening test for PA (see below); (2) a
93 diagnosis of confirmed or not confirmed PA by confirmatory testing. Patients were excluded in case
94 of autonomous cortisol secretion. Eligible patients from the developmental cohort were randomized
95 to a training cohort (n=522) or to an internal validation cohort (n=174). An independent prospective
96 cohort of 328 patients consecutively recruited from the Munich Klinikum der Universität was used
97 for external validation. All the patients included in the present retrospective analysis gave extended
98 written consent for the use of their personal data, according to Helsinki declaration. The study was
99 approved by local ethical committees.

100 PA was diagnosed in accordance with the Endocrine Society (ES) guideline [1] and ESH consensus
101 [2,6]. Screening was performed by measurement of the aldosterone concentration (AC) to plasma
102 renin activity ratio (ARR) in the developmental cohort and by AC to direct renin concentration (DRC)
103 in the external validation cohort. Interfering drugs were withdrawn according to guidelines [1,2]. The

104 screening test was considered positive if ARR was higher than 30 ng/dL/ng*mL⁻¹*h⁻¹ and AC higher
105 than 10 ng/dL. Patients with suspected PA underwent confirmatory testing by either an intravenous
106 saline loading test or a captopril challenge test [2]. The cut-offs for a positive confirmatory test were
107 a post-test AC > 5 ng/dL for intravenous saline loading, or a post-test ARR > 30 ng/dL/ng*mL⁻¹*h⁻¹
108 for a captopril challenge test. After PA confirmation, subtype diagnosis was defined by computed
109 tomography and adrenal venous sampling [6].

110

111 *Statistics and machine learning analyses*

112 Kolmogorov–Smirnov test was used to evaluate the distribution of patient parameters. Normally
113 distributed parameters were expressed as mean ± standard deviation and analyzed by student t-test.
114 Non-normally distributed parameters were expressed as median [interquartile range] and analyzed by
115 Mann-Whitney’s test. Categorical parameters were expressed as absolute number and percentage
116 distribution and analyzed by Chi-square test. A *P*-value of less than 0.05 was considered significant.
117 Univariate and multivariate logistic regression was used to assess the odds ratios (ORs). An OR
118 greater than 1 was associated with an increased likelihood of a confirmed diagnosis of PA, an OR
119 less than 1, a decreased likelihood.

120 The machine learning models, and the PACT score were built on the training cohort and then tested
121 in the internal and external validation cohorts. Patients from the internal and external validation
122 cohorts were distinct from the training cohort, in which all the models were built.

123 Supervised machine learning algorithms are used to formulate prediction about a select outcome on
124 the base of a pre-defined set of labeled paired input-output data [7]. We used different models,
125 including linear discriminant analysis (LDA), random forest (RF) classification algorithms, and
126 support vector machine (SVM) with different kernels (linear and gaussian radial basis function). LDA
127 employs linear combination of parameters to maximize the separation between groups by increasing
128 precision estimates by variance reduction. The predicted diagnosis is derived from the following
129 equation: Confirmed PA diagnosis = LDAcoeff₁*Variable₁ + LDAcoeff₂*Variable₂ + ... +

130 LDAcoeff_n*Variable_n > tested thresholds. The RF algorithm creates 30 classification trees with a
131 maximum number of 7 splits for each tree. The predicted diagnosis resulted from the outcome of each
132 classification tree of the forest; if at least 16 of 30 trees of the RF confirm PA, then the diagnosis of
133 PA will be confirmed. Linear SVM builds a classification model to assign patients to their diagnosis
134 given a linear boundary. The model finds out the plane which best separates groups of patients (i.e.
135 confirmed vs. not confirmed diagnosis of PA), maximizing the distances between them. Patients are
136 classified according to the following equation: SVMcoeff₀ + SVMcoeff₁*Variable1 +
137 SVMcoeff₂*variable₂ + + SVMcoeff_n*Variable_n. Gaussian SVM allows to divide patients using
138 a non-linear boundary. The corresponding equation is: SVMcoeff₀ + SVMcoeff₁*f(Variable1) +
139 SVMcoeff₂*f(variable₂) + + SVMcoeff_n*f(Variable_n), where “f” is an exponential function
140 coefficient.

141 The diagnostic performance of the PACT score was assessed by analysis of receiver operating
142 characteristics (ROC) curves. The area under the curve (AUC) was evaluated to define the best cut-
143 off by Youden Index (J = sensitivity + specificity - 1). Overfitting effect was defined as the difference
144 between the accuracy at the training of the models and the accuracy at validation. A free-
145 downloadable tool was developed to calculate the score and the predicted diagnosis (available at:
146 <https://github.com/CentroIpertenUnito/PACT-score/raw/master/PACT%20Score%20Calculator.xlsm>).

147 Python 3.5 (library, scikit-learn) and IBM SPSS Statistics 26 (IBM Corp., Armonk, New York, USA)
148 were used for analysis.

149

150 **RESULTS**

151 *Clinical and biochemical characteristics*

152 In the present study, we evaluated the data from 1,024 patients including a developmental cohort
153 from Torino (n=696) and a prospective independent validation cohort from Munich (n=328). Patients
154 from the developmental cohort included 421 subjects with a confirmed diagnosis of PA, and 275 for
155 whom PA was not confirmed (Table 1). Patients with confirmed PA were more frequently male, with

156 higher systolic blood pressure, and defined daily dose (DDD) for antihypertensive medications,
157 compared with patients in whom PA was not confirmed ($p<0.01$ for all comparisons). Aldosterone
158 levels were higher, and plasma renin activity and potassium levels were lower in patients with
159 confirmed vs. not confirmed PA ($p<0.001$). The prevalence of organ damage and cardiovascular
160 events was higher in patients with a confirmed diagnosis of PA.

161 Univariate logistic regression analysis confirmed a relevant association of sex, systolic blood
162 pressure, DDD, PRA and aldosterone levels at screening, lowest potassium, and prevalence of organ
163 damage and cardiovascular events with a confirmed diagnosis of PA (Table S1). Multivariate
164 regression analysis confirmed female sex (OR 0.41), DDD (OR 1.20), PRA (OR 0.07), aldosterone
165 (OR 1.08), lowest potassium (OR 0.14), and presence of organ damage (OR 2.63) as independent
166 predictors of a confirmed diagnosis of PA (Table 2 and Table S2).

167

168 ***Diagnostic modelling***

169 Patients from the developmental cohort were randomly assigned to a training cohort (n=522) and to
170 an internal validation cohort (n=174). No differences were found between the 2 cohorts with regard
171 to clinical and biochemical parameters (Table S3). All diagnostic models were developed in the
172 training cohort and tested in the internal validation cohort. A second external cohort of patients was
173 used for an independent validation (n=328). Patients from the external validation cohort compared
174 with the developmental cohort, displayed lower aldosterone levels, lower potassium, lower DDD,
175 shorter duration of hypertension, and lower systolic blood pressure ($p<0.01$ for all comparisons). The
176 prevalence of PA was also lower than in the developmental cohort (52.7 vs. 60.5%; $p=0.019$; Table
177 S4).

178 The 6 parameters selected by regression analysis (Table 2) were used for the development of machine
179 learning models. Their diagnostic performance in the discrimination of patients with a confirmed
180 diagnosis of PA on the combined developmental cohort is shown in Figure S1. Predictor importance
181 coefficients are reported in Figure S2. The best predictor was lowest potassium in all the models.

182 The linear combination of parameters by LDA is shown in the canonical plot (Figure S1A). In
183 particular, 551 of 696 patients (accuracy 79.2%) were correctly classified, with a sensitivity and
184 specificity for PA detection of 84.1% and 71.6%, respectively.

185 The RF classification algorithm reached a higher performance by correctly discriminating 571 of 696
186 patients (accuracy 82%) with a sensitivity and specificity of 87.9% and 73.1%, respectively. The RF
187 was composed by 30 trees; the first tree of the series is reported in Figure S1B.

188 Finally, SVM models (linear and gaussian kernel) displayed similar performance, with an accuracy
189 of 80.0% and 81.6%, respectively. The linear SVM was able to correctly classify 354 of 421 patients
190 with a confirmed diagnosis of PA (sensitivity 84.1%) and 203 of 275 patients with not confirmed PA
191 (specificity 73.8%). The gaussian SVM correctly classified 365 of 421 patients with a confirmed
192 diagnosis of PA (sensitivity 86.7%), at the same specificity of the linear kernel (73.8%).

193 Representative plots for main discriminants of SVM models are shown in Figure S1C and S1D.

194 Table S5 reports the confusion matrix and diagnostic performance of machine learning based models
195 at training and internal validation on the developmental cohort. The overfitting effect was low for all
196 the models (between 2.1% and 9.2%), thus suggesting an acceptable generalizability of the models.
197 The performance at external validation on the independent cohort was still very high, ranging between
198 72.9% and 78.7% (Table S6).

199

200 ***Development and validation of the PACT score***

201 The same 6 parameters employed for the machine learning models were used to develop a 16-point
202 scoring system. As for the diagnostic algorithms described above, the PACT score was built on the
203 training cohort and then tested in the internal and external validation cohorts (Table S7).
204 Categorization and points assignment are shown in Figure 1A and 1E.

205 The PACT score was directly correlated with the proportion of patients with a confirmed diagnosis
206 of PA (Figure 1B) and with pre-/post- confirmatory testing aldosterone and aldosterone-to-renin ratio
207 (Pearson's R ranging between 0.247 and 0.479; $p < 0.001$; Table S8). Noteworthy, for all patients with

208 a score greater than 12, PA diagnosis was confirmed, whereas for all patients with a score lower than
209 5, PA diagnosis was not confirmed (Table S9). The analysis of ROC curves demonstrated a reliable
210 performance of the score both at screening and internal validation (AUC of 0.879 and 0.877,
211 respectively; Figure 1C and 1D). The cut-off with the highest accuracy was 8 points.

212 In the combined developmental cohort, a score equal or greater than 8 correctly discriminated a
213 confirmed diagnosis of PA in 388 of 421 patients (sensitivity 92.2%), whereas a score lower than 8
214 identified those with not confirmed PA in 197 of 275 cases (specificity 71.6%), with an overall
215 accuracy of 84.1%. A cut-off of equal or greater than 5 reached 100% sensitivity, correctly classifying
216 all the patients with a confirmed diagnosis of PA, both in the training and validation cohorts. On the
217 other side, a score lower than 13 correctly identified all the patients for which a diagnosis of PA was
218 not confirmed at training and at validation of the score. Confusion matrix and diagnostic performance
219 of the PACT score at training and internal validation are reported in Table S7.

220 Accuracy at internal validation was 83.9%. The overfitting effect was minimum (0.2%), and
221 consistently the performance at external validation confirmed the very high generalizability of our
222 score system with an accuracy of 81.1% (Table S6). The PACT score displayed a sensitivity ranging
223 between 78.6% and 91.9% and specificity between 73.3% and 83.9%, at internal and external
224 validation. Of note, this performance was similar and even higher than that of machine learning
225 models (accuracy at external validation: 78.4%, 72.9%, 78.4%, and 78.7%, for the LDA, the RF
226 algorithm, linear and gaussian SVM, respectively).

227

228 ***Management of patients with PA***

229 The PACT score was implemented in a flow chart for the management of patients with PA (Figure
230 2). Patients with a positive screening test were stratified for their likelihood of PA diagnosis according
231 to our score-system (developmental cohort and external validation cohort; n=1,024). For patients with
232 a score less than 5, PA diagnosis was excluded without a confirmation test (n=107); instead, for
233 patients with a score equal or greater than 13, PA diagnosis was confirmed without further tests

234 (n=126). All the remaining patients (n=791; score between 5 and 12) should undergo confirmatory
235 test and be allocated according with subsequent investigations. This approach resulted in the correct
236 management of all patients (accuracy 100%), with the reduction of 22.8% of unnecessary
237 confirmatory tests (233 of 1,024 procedures).

238 A second model for the management of patients with PA was developed by the use of stricter cut-
239 offs (Figure S3A). In this case, a PACT score lower than 6 correctly excluded PA diagnosis in 170
240 of 191 patients; 21 patients with a confirmed diagnosis of PA were missed (none of them displayed
241 a diagnosis of aldosterone producing adenoma). A cut off of equal or greater than 11 correctly
242 confirmed the diagnosis of PA in 258 of 277 patients; 19 patients with low-renin hypertension would
243 undergo to inappropriate adrenal venous sampling. At this regard, we would underline that for 11 of
244 the 19 misclassified low-renin hypertensive patients, the confirmatory test was performed before
245 2014 in recumbent position, thus representing potentially false negative patients [8]. The second flow
246 chart displayed an accuracy of 96.1% (96.5% and 95.6% of sensitivity and specificity, respectively),
247 reducing the number of necessary confirmatory tests of 45.7% (Figure S3B).

248 Finally, we performed a sub-analysis on patients with unilateral PA. The performance of all the
249 proposed diagnostic models is reported in Table S10: the accuracy ranged between 95 and 100%. In
250 particular, the two flow charts for patient management correctly classified as confirmed PA all the
251 patients with unilateral disease.

252

253 **DISCUSSION**

254 We used supervised machine-learning algorithms and regression analysis to develop prediction
255 models and a simple scoring system, to discriminate patients with confirmed PA. We provided
256 evidence that confirmatory testing could be avoided in a subset of patients selected by their clinical
257 and biochemical parameters at screening. We developed a flow-chart integrating the PACT score to
258 stratify patients with suspected PA and a positive screening test according to their likelihood of PA
259 diagnosis. This model provides options to clinicians, who may propose a strict follow-up for patients

260 with a low likelihood of PA (PACT score less than 5), or directly proceed to subtype diagnosis with
261 CT scanning and adrenal venous sampling for patients with a high likelihood of PA (PACT score
262 equal or greater than 13). Patients with an intermediate risk should follow the diagnostic flow-chart
263 recommended by the guidelines and undergo confirmatory testing. This approach results in the correct
264 management of all patients and reduces the number of confirmatory tests by 23%. An online
265 downloadable tool facilitates the application of the PACT score in clinical practice.

266

267 The growing awareness of the scientific community on the importance of screening for secondary
268 causes of hypertension [9], will lead to an increase in the number of patients with a positive screening
269 test for PA, and therefore an increased requirement of confirmatory tests.

270 The systematic confirmation of all patients with a positive screening test determines an increase in
271 costs, time, risks and complexity in the management of patients with PA [10], thus contributing to
272 the underdiagnosis of PA [11]. The PACT score may simplify the diagnostic work-up for a subset of
273 selected patients at low or high likelihood of PA, resulting in an increased availability of resources to
274 be allocated for subtype diagnosis, and targeted therapy, which are efficacious and cost-effective [12].
275 According to the ES guideline and ESH consensus, patients with hypokalemia with suppressed renin
276 and high aldosterone levels could skip confirmatory testing [1,2]; in the overall population included
277 in the study (developmental and validation cohorts), 45 patients (4.4%) displayed these characteristics
278 and could directly undergo subtype diagnosis. On the other side, 126 patients (12.3%) could directly
279 undergo subtype diagnosis following the PACT score, and in another 10.4% of cases, the diagnosis
280 of PA could be excluded without further testing, resulting in a net advantage on confirmatory testing
281 reduction (23% with the PACT score, compared to 4.4% with the ES recommendations). Noteworthy,
282 using stricter cut-offs the PACT score could allow the reduction of 46% of confirmatory tests
283 maintaining an overall accuracy of 96.1%.

284 Avoiding potentially unnecessary confirmatory tests could have an impact not only on the reduction
285 of costs, but also on patient management. Even if minimally invasive, confirmatory testing is

286 associated with side effects, including hypertensive or hypotensive episodes, arrhythmias, vertigo,
287 headache, dyspnea, and neurological symptoms [10]. A reduction of the number of tests will also
288 reduce the incidence of clinical complications, related to intravascular volume expansion (in case of
289 saline loading test) or hypotension (after captopril administration).

290

291 Previous studies proposed different cut-offs for AC and ARR to avoid confirmatory tests. Nanba et
292 al. first demonstrated that most patients evaluated with saline infusion test, captopril challenge test
293 and/or furosemide upright test displayed a confirmed diagnosis of PA when ARR was equal or greater
294 than 100 [ng/dL]/[ng/mL/h] in the presence of an AC of at least 25 ng/dL [13]. Maiolino et al.
295 observed that the progressive increase of the ARR at screening was associated with an increased
296 specificity for the diagnosis of an aldosterone producing adenoma [14]. Recommendations from the
297 French Endocrinology and Hypertension societies suggest the avoidance of confirmatory testing in
298 the presence of an AC above 20 ng/dL, high ARR, with or without hypokalemia. On the contrary, PA
299 diagnosis could be ruled out if AC at screening is less than 9 ng/dL in two different occasions [15].
300 This approach resulted in a positive predictive value of 93% in a recent study on a cohort of 173
301 hypertensive patients referred to a single hypertension center [16]. Finally, two recent reports
302 observed that an AC above 30 ng/dL, or above 20 ng/dL in the presence of hypokalemia, rendered
303 confirmatory testing unnecessary [17,18]. Main limitations of all these studies are represented by the
304 general applicability of the proposed cut-offs, the absence of an independent validation, and the
305 relatively low sensitivity and negative predictive value.

306 To our knowledge, one single study combined different clinical parameters to develop a scoring
307 system to skip confirmatory testing [19]. Using age, BMI, number of antihypertensive medications,
308 sodium, potassium, and presence of diabetes, the authors reported 100% specificity and positive
309 predictive value for PA diagnosis, with a reduction of 42.2% of confirmatory test. However, this was
310 a single-center retrospective study with a limited sensitivity (52%) not validated in an independent
311 cohort [19]. In the PACT score, we combined parameters which are easily available for patients who

312 underwent screening for PA: sex, DDD, PRA and aldosterone values, potassium, and the presence of
313 organ damage. In line with previous reports, hypokalemia, aldosterone and PRA at screening were
314 the main discriminants [17,18,20], with lowest potassium being the best predictor in all machine
315 learning models. In particular, an AC of at least 28 ng/dL and a PRA less than 0.2 mg/mL/h, in
316 presence of history of hypokalemia (lowest potassium < 3.7 mEq/L) resulted in a score of 11,
317 corresponding to a 77.8% likelihood of PA. Male patients had an increase probability of a confirmed
318 diagnosis of PA and female sex has been associated with a false positive result at screening test
319 [21,22]. Antihypertensive medications expressed as DDD and target organ damage completed the
320 score.

321

322 Some study limitations should be acknowledged. First of all, the generalizability of our approach for
323 patient management could be limited by bias related to the cohort characteristics and assays used for
324 AC and PRA measurement. However, we tested our diagnostic models in an external validation
325 cohort which differed significantly for 5 of the 6 included parameters, and where patients were
326 screened using DRC instead of PRA and displayed lower median aldosterone levels. Moreover, we
327 built our score in a retrospective developmental cohort, but the validation was provided in an
328 independent external prospective cohort of patients enrolled consecutively. Finally, the PACT score
329 has been designed to select patients at screening test to avoid a proportion of confirmatory tests, but
330 it was not tested in patients who had PRA and aldosterone measured under interfering medications.
331 In this regard, it should be noted that the effects of interfering drugs on hormonal measurements
332 cannot be standardized, which is a major issue when developing a prediction model. The assessment
333 of the diagnostic performance of our models in a cohort of patients screened for PA under interfering
334 drugs should be evaluated in future studies.

335

336 This is the first study which implemented machine learning algorithms and regression analysis to
337 develop and validate prediction models to discriminate patients with a confirmed diagnosis of PA,

338 using screening parameters, in a large cohort of patients from two specialized referral centers.
339 Internal and external validation of the models demonstrated a reliable generalizability with a low
340 overfitting effect. The performance of the proposed diagnostic algorithms was higher than all
341 previously proposed approaches, with a reduction of the number of confirmatory tests of up to 23%.
342 The algorithm appears to be a significant improvement compared with recommendations of
343 international guidelines, which would have bypassed confirmatory testing in 4.4% of patients. This
344 approach may have a high potential impact on the management of PA with a reduction of costs and
345 simplification of the diagnostic work-up of patients with hypertension.

346

347 **CONCLUSIONS**

348 Combining different clinical parameters in the PACT score, we discriminate with high accuracy
349 patients with a confirmed diagnosis of PA, in a large cohort of patients with a positive screening test.
350 This approach could result in a significant reduction of unnecessary confirmatory tests. The
351 integration of diagnostic modelling algorithms in clinical practice will increase the detection rate of
352 PA and improve the management of these patients.

353

354 **Source(s) of Funding:** This research did not receive any specific grant in Torino. The German Conn-
355 Registry-Else-Kröner Hyperaldosteronism Registry is supported by the Else Kröner-Fresenius
356 Stiftung (2013_A182, 2015_A171 and 2019_A104). CA and MR are supported by the Deutsche
357 Forschungsgemeinschaft (DFG, within the CRC/Transregio 205/1 “The Adrenal: Central Relay in
358 Health and Disease”).

359

360 **Conflict(s) of Interest/Disclosure(s):** the authors have nothing to disclose.

361

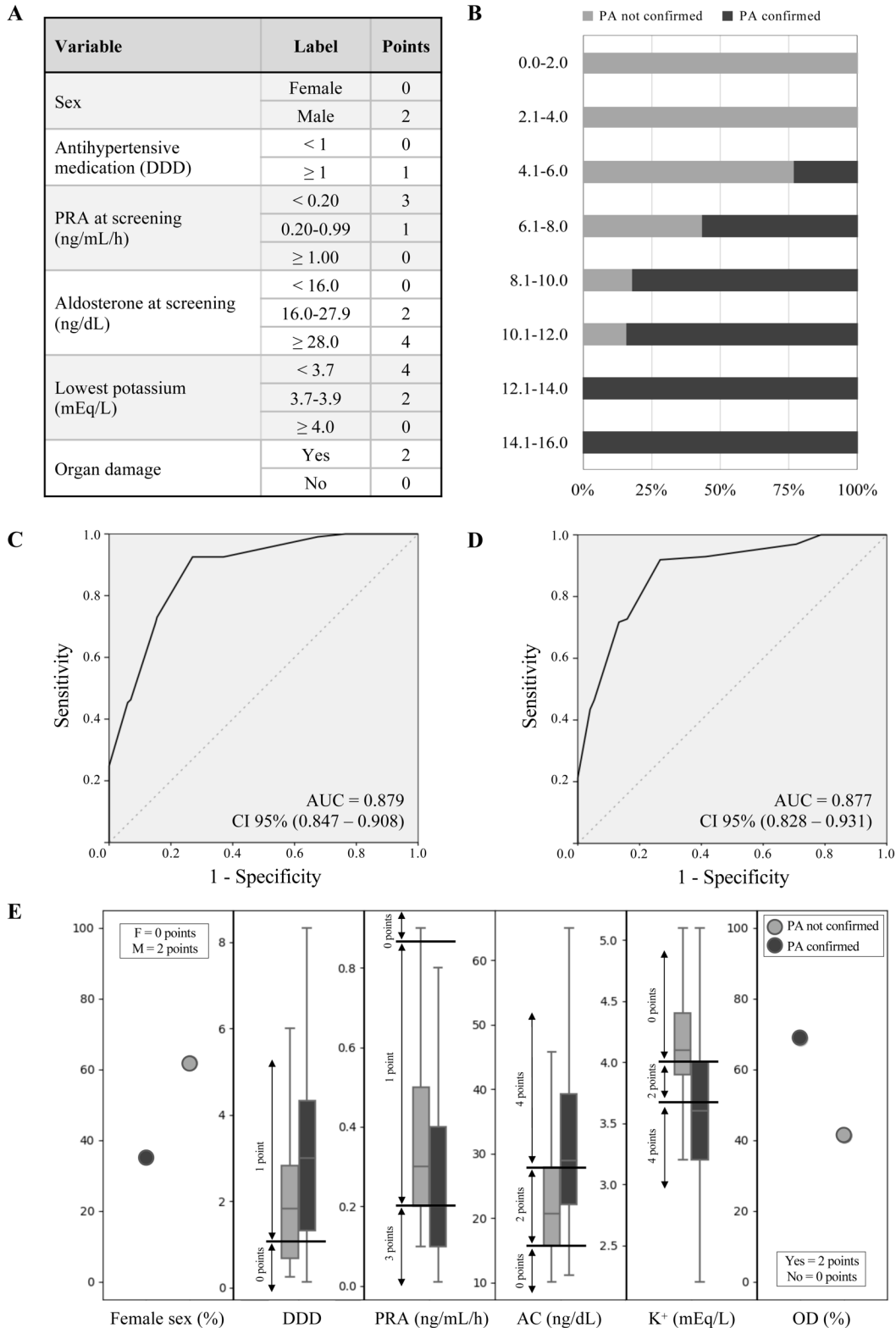
362

363 **REFERENCES**

- 364 1. Funder JW, Carey RM, Mantero F, et al. The Management of Primary Aldosteronism: Case
365 Detection, Diagnosis, and Treatment: An ES Clinical Practice Guideline. *J Clin Endocrinol*
366 *Metab.* 2016;101:1889-916.
- 367 2. Mulatero P, Monticone S, Deinum J, et al. Genetics, prevalence, screening and confirmation of
368 primary aldosteronism: a position statement and consensus of the Working Group on Endocrine
369 Hypertension of The European Society of Hypertension. *J Hypertens* 2020;38:1919-28.
- 370 3. Monticone S, Burrello J, Tizzani D, et al. Prevalence and Clinical Manifestations of Primary
371 Aldosteronism Encountered in Primary Care Practice. *J Am Coll Cardiol.* 2017;69:1811-20.
- 372 4. Nishikawa T, Omura M, Satoh F, et al; Task Force Committee on Primary Aldosteronism, The
373 Japan Endocrine Society. Guidelines for the diagnosis and treatment of primary aldosteronism--
374 the Japan Endocrine Society 2009. *Endocr J.* 2011;58:711-21.
- 375 5. Burrello J, Amongero A, Buffolo F, et al. Data from: Development and validation of a prediction
376 score to avoid confirmatory testing in patients with suspected primary aldosteronism. *J Clin*
377 *Endocrinol Metab* 2020. <https://github.com/CentroIpertenUnito/PACT-score>. Deposited 15
378 December 2020.
- 379 6. Mulatero P, Sechi LA, Williams TA, et al. Subtype diagnosis, treatment, complications and
380 outcomes of primary aldosteronism and future direction of research: a position statement and
381 consensus of the Working Group on Endocrine Hypertension of the European Society of
382 Hypertension. *J Hypertens.* 2020;38:1929-36.
- 383 7. Burrello J, Burrello A, Stowasser M, et al. The Primary Aldosteronism Surgical Outcome Score
384 for the Prediction of Clinical Outcomes After Adrenalectomy for Unilateral Primary
385 Aldosteronism. *Ann Surg.* 2020;272:1125-32.
- 386 8. Ahmed AH, Cowley D, Wolley M, et al. Seated saline suppression testing for the diagnosis of
387 primary aldosteronism: a preliminary study. *J Clin Endocrinol Metab.* 2014;99:2745-53.

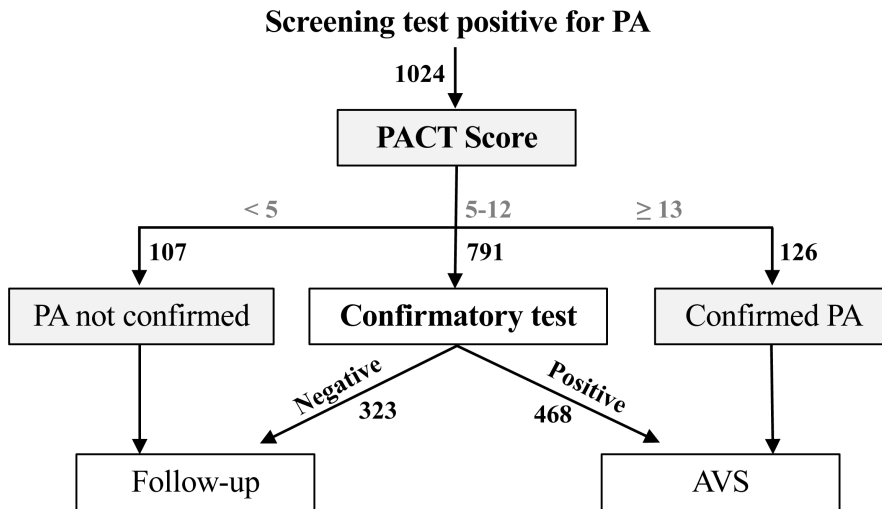
- 388 9. Olsen MH, Angell SY, Asma S, et al. A call to action and a lifecourse strategy to address the
389 global burden of raised blood pressure on current and future generations: the Lancet Commission
390 on hypertension. *Lancet*. 2016;388:2665-12.
- 391 10. Heinrich DA, Adolf C, Quinkler M, et al. Safety of medical adjustment and confirmatory testing
392 in the diagnostic work-up of primary aldosteronism. *Eur J Endocrinol*. 2019;181:421-8.
- 393 11. Mulatero P, Monticone S, Burrello J, Veglio F, Williams TA, Funder J. Guidelines for primary
394 aldosteronism: uptake by primary care physicians in Europe. *J Hypertens*. 2016;34:2253-7.
- 395 12. Lubitz CC, Economopoulos KP, Sy S, et al. Cost-Effectiveness of Screening for Primary
396 Aldosteronism and Subtype Diagnosis in the Resistant Hypertensive Patients. *Circ Cardiovasc*
397 *Qual Outcomes*. 2015;8:621-30.
- 398 13. Nanba K, Tamanaha T, Nakao K, et al. Confirmatory testing in primary aldosteronism. *J Clin*
399 *Endocrinol Metab*. 2012;97:1688-94.
- 400 14. Maiolino G, Rossitto G, Bisogni V, et al; PAPY Study Investigators. Quantitative Value of
401 Aldosterone-Renin Ratio for Detection of Aldosterone-Producing Adenoma: The Aldosterone-
402 Renin Ratio for Primary Aldosteronism (AQUARR) Study. *J Am Heart Assoc*. 2017;6:e005574.
- 403 15. Reznik Y, Amar L, Tabarin A. SFE/SFHTA/AFCE consensus on primary aldosteronism, part 3:
404 Confirmatory testing. *Ann Endocrinol (Paris)*. 2016;77:202-7.
- 405 16. Vivien M, Deberles E, Morello R, Haddouche A, Guenet D, Reznik Y. Evaluation of
406 Biochemical Conditions Allowing Bypass of Confirmatory Testing in The Workup of Primary
407 Aldosteronism: A Retrospective Study in a French Hypertensive Population. *Horm Metab Res*.
408 2019;51:172-177.
- 409 17. Umakoshi H, Sakamoto R, Matsuda Y, et al. Role of Aldosterone and Potassium Levels in
410 Sparing Confirmatory Tests in Primary Aldosteronism. *J Clin Endocrinol Metab*.
411 2020;105:dgz148.

- 412 18. Kawashima J, Araki E, Naruse M, et al. Baseline Plasma Aldosterone Level and Renin Activity
413 Allowing Omission of Confirmatory Testing in Primary Aldosteronism. *J Clin Endocrinol*
414 *Metab.* 2020;105:dgaal117.
- 415 19. Kietsiroje N, Wonghirundecha R, Suntornlohanakul O, Murray RD. Construction of a
416 predictive scoring system as a guide to screening and confirmation of the diagnosis of primary
417 aldosteronism. *Clin Endocrinol (Oxf)*. 2020;92:196-205.
- 418 20. Burrello J, Monticone S, Losano I, et al. Prevalence of Hypokalemia and Primary Aldosteronism
419 in 5100 Patients Referred to a Tertiary Hypertension Unit. *Hypertension*. 2020;75:1025-33.
- 420 21. Pizzolo F, Raffaelli R, Memmo A, et al. Effects of female sex hormones and contraceptive pill
421 on the diagnostic work-up for primary aldosteronism. *J Hypertens*. 2010;28:135-42.
- 422 22. Ahmed AH, Gordon RD, Taylor PJ, Ward G, Pimenta E, Stowasser M. Are women more at risk
423 of false-positive primary aldosteronism screening and unnecessary suppression testing than men?
424 *J Clin Endocrinol Metab*. 2011;96:E340-6.



425
 426 **Legend to Figure 1 – Development of the PACT score.** Univariate/multivariate regression analyses
 427 were used to assign points to each variable according to stratification level. The score was developed
 428 in the training cohort (n=522) and tested on the internal validation cohort from Torino (n=174). Data
 429 on training and validation of the score are reported in Table S7. (A) The table reports included
 430 variables and scoring-point system. If only direct renin concentration (DRC) is available, the
 431 following cut-offs could be used: DRC < 2.5 mU/L (3 points); DRC 2.5-12.3 mU/L (1 point); DRC
 432 ≥ 12.4 mU/L (0 point). (B) Histogram showing the proportion of patients (x-axis, %) for each
 433 diagnosis (PA confirmed, black; PA not confirmed, grey), stratified by score points (y-axis) on the

434 developmental cohort (n=696). The total number of patients (N) for each score level and their
 435 proportion (%) are reported in Table S9. (C, D) Receiver operating characteristics (ROC) curve to
 436 assess the area under the curve (AUC) in the training (n=522; left) and internal validation cohort from
 437 Torino (n=174; right). (E) Representation of variable categorization and assigned points (PA
 438 confirmed, black; PA not confirmed, grey); the bars indicate median and interquartile range.
 439
 440



441
 442 *Legend to Figure 2 – Management of patients with suspected PA.* Flow chart for the management
 443 of patients with a positive screening test according to the PACT score (Developmental Cohort +
 444 External Validation Cohort; n=1,024). The number of patients is indicated in bold; cut-offs are
 445 indicated in grey. AVS, Adrenal Venous Sampling; PA, Primary Aldosteronism; PACT, Primary
 446 Aldosteronism Confirmatory Testing Score.

447 **Table 1. Patient Characteristics**

Variable	Confirmatory / Exclusion Test		P-value
	PA confirmed (n=421)	PA not confirmed (n=275)	
Age at diagnosis (years)	50 ± 10.2	51 ± 9.5	0.202
Female sex, n (%)	148 (35.2)	170 (61.8)	<0.001
Duration of HTN (months)	68 [24; 135]	54 [18; 125]	0.084
Systolic BP (mmHg)	157 ± 20.7	152 ± 19.4	0.003
Diastolic BP (mmHg)	95 ± 11.2	94 ± 10.5	0.051
Antihypertensive medication (DDD)	3.00 [1.33; 4.33]	1.83 [0.67; 2.83]	<0.001
BMI (Kg/sqm)	25.9 ± 4.48	25.5 ± 3.95	0.204
PRA at screening (ng/mL/h)	0.20 [0.10; 0.40]	0.30 [0.20; 0.50]	<0.001
Aldosterone at screening (ng/dL)	28.9 [22.1; 39.3]	20.7 [15.7; 27.9]	<0.001
Lowest Potassium (mEq/L)	3.6 ± 0.63	4.1 ± 0.41	<0.001
eGFR (mL/min)	91 ± 17.1	91 ± 16.9	0.666
Diabetes, n (%)	31 (7.4)	13 (4.7)	0.162
Organ damage, n (%)	290 (68.9)	114 (41.5)	<0.001
CV events, n (%)	61 (14.5)	20 (7.3)	0.004

448

449 Characteristics of patients included in the analysis: confirmed diagnosis of PA (n=421) vs. PA not
450 confirmed (n=275). HTN, Hypertension; BP, Blood Pressure; DDD, Defined Daily Dose (average
451 maintenance dose per day for a drug used for its main indication in adults); PRA, Plasma Renin
452 Activity; eGFR, estimated Glomerular Filtration Rate; CV, Cardiovascular. Organ damage is defined
453 as presence of left ventricular hypertrophy at echocardiography and/or microalbuminuria. Normally
454 and non-normally distributed variables were reported as mean ± standard deviation or median
455 [interquartile range], respectively. Categorical variables were reported as absolute number (n) and
456 proportion (%).

457 **Table 2. Regression analysis on discriminant for PA diagnosis**

Variable (ref. PA confirmed)	Univariate analysis		Multivariate Analysis	
	OR (CI 95%)	<i>P</i> -value	OR (CI 95%)	<i>P</i> -value
Female sex, n (%)	0.34 (0.24-0.46)	<0.001	0.41 (0.28-0.62)	<0.001
Antihypertensive medication (DDD)	1.40 (1.27-1.54)	<0.001	1.20 (1.07-1.35)	0.002
PRA at screening (ng/mL/h)	0.28 (0.14-0.57)	<0.001	0.07 (0.03-0.20)	<0.001
Aldosterone at screening (ng/dL)	1.06 (1.04-1.08)	<0.001	1.08 (1.06-1.10)	<0.001
Lowest Potassium (mEq/L)	0.13 (0.09-0.19)	<0.001	0.14 (0.09-0.23)	<0.001
Organ damage, n (%)	3.13 (2.28-4.29)	<0.001	2.63 (1.75-3.95)	<0.001

458

459 Odds ratio (OR) and the 95% confidence interval (CI) were evaluated by univariate and multivariate
 460 logistic regression analysis, as indicated. An OR greater than 1 indicates an increased likelihood of
 461 confirmed PA, and an OR less than 1 a decreased likelihood (i.e., an OR of 1.06 for aldosterone levels
 462 means an increase of 6% in the likelihood of confirmed PA, for each 1 ng/dL increase of aldosterone;
 463 an OR of 0.07 for PRA means an increase of 43% in the likelihood of confirmed PA, for each 0.1
 464 ng/mL/h decrease in PRA). Antihypertensive medication (expressed as DDD), PRA and aldosterone
 465 at screening, and lowest potassium were treated as continuous variables; sex and organ damage were
 466 treated as categorical variables.