Archives of Cardiovascular Disease xxx (xxxx) xxx–xxx

Disponible en ligne sur
ScienceDirect
www.sciencedirect.com

Elsevier Masson France
EM|consulte
www.em-consulte.com

Clinical Research

Electrocardiographic phenotypes of a representative subset of the French general population: ECGs at inclusion in the CONSTANCES cohort

Fabrice Extramiana[a,b,*], Thien Vu Nguyen Minh Tuan[a], Martino Vaglio[c], Anna Ozguler[d], Peter W. Macfarlane[e], Pierre Maison-Blanche[a], Fabio Badilini[c]

[a] AP–HP, Unité Rythmologie, Hôpital Bichat, 75018 Paris, France
[b] Université Paris Cité, 75006 Paris, France
[c] AMPS LLC, NY, NY 10025, USA
[d] UMS11 Inserm, 'Cohortes en Population', 94800 Villejuif, France
[e] College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow G12 8QQ, UK

ARTICLE INFO

Keywords:
Electrocardiogram
Automated analyses
Epidemiology

ABSTRACT

Background: Large electrocardiogram (ECG) dataset analyses have emerged as potential game-changers in the field of personalized predictive medicine. ECG parameters have been described in cohorts of apparently healthy subjects and from primary care but seldom in community-based representative populations.
Aims: To describe the ECG phenotypes of a representative subset of the adult French general population.
Methods: ECGs recorded at inclusion in the CONSTANCES cohort were automatically analysed using the Glasgow diagnostic algorithm. Extreme values and abnormal statements were adjudicated to detect false positives. A subset of ECGs that were classified as normal were also adjudicated to estimate false negative statements. The data obtained were used to describe the prevalence and distribution of quantitative parameters and diagnostic statements.
Results: We automatically analysed the ECGs of 143,763 subjects (54% female; mean ± standard deviation age 47.0 ± 13.5 years in females and 46.9 ± 13.5 years in males; $P = 0.44$) and adjudicated > 10,000 ECGs. We describe the distribution of automatic ECG interval measurements and the prevalence of different ECG statements provided by the automatic analysis, before and after adjudication. Heart rate and interval durations were dependent on both sex and age (ANOVA $P < 0.0001$). At the population level, the Fridericia formula appeared to be less biased than that of Bazett.
Conclusions: We describe the adjudicated ECG phenotypes of a representative subset of the adult French general population. The automated measurement and interpretation provided by the Glasgow algorithm proved highly efficient for epidemiological evaluation. This ECG phenotype characterization will open up the possibilities for cross-analyses, for artificial intelligence-based algorithm development and will serve as a reference to evaluate serial ECG changes during extended follow-up.

1. Background

Cardiovascular diseases (including ischaemic heart disease and stroke) remain the leading causes of age-standardized deaths globally [1]. Despite important improvements over the last 3 decades [1], prevention and prediction of cardiovascular diseases continue to present a global public health priority [2]. Even in the era of multimodal cardiac imaging, the electrocardiogram (ECG) is the most commonly conducted cardiovascular diagnostic procedure. The ECG is the cornerstone for the diagnosis of ischaemic and other structural heart diseases and is essential for the diagnosis of arrhythmias including atrial fibrillation [3].

At the population level, the United States Preventive Service Task Force [4,5] does not recommend ECG screening for risk of cardiovascular disease or atrial fibrillation. However, digitally based quantitative electrocardiology [6–8], serial ECG change analyses [9] and, more recently, artificial intelligence-based ECG analyses [10–13] have emerged

Abbreviations: ANOVA, analysis of variance; ECG, electrocardiogram; LBBB, left bundle branch block; RBBB, right bundle branch block; QTc, corrected QT; QTcB, Bazett's corrected QT interval; QTcF, Fridericia's corrected QT interval; SD, standard deviation.

* Corresponding author. Cardiology department, Bichat hospital, 46, Rue Henri-Huchard, 75018 Paris, France.
E-mail address: fabrice.extramiana@aphp.fr (F. Extramiana).

https://doi.org/10.1016/j.acvd.2025.12.013
Received 1 July 2025; Received in revised form 24 December 2025 Accepted 29 December 2025
Available online xxx

as potential game-changers in the field of personalized predictive medicine.

These developments require collection of large ECG databases. Many studies have described ECG databases from large cohorts of 'normal' or 'apparently healthy' subjects from diverse origins [14–23] and from primary care or pharmaceutical company-sponsored clinical trials [24–27]. Conversely, ECG data from community-based representative populations have only seldom been described [28–34].

The CONSTANCES cohort is a large French population cohort (n = 220,000 subjects) selected to be representative of the adult French population aged 18–69 years at selection [35,36]. Due to a delay between selection and first ECG, a small subset of subjects were older than 70 years. CONSTANCES aims at better understanding the determinant of diseases using a multifactorial approach (biological, genetic, but also environmental and sociological). Subjects included in the CONSTANCES cohort undergo questionnaires and a large batch of tests, including an ECG, at inclusion and every 4 years thereafter. These ECGs will serve as a massive longitudinal digital ECG database, opening the possibility for automated analyses, serial ECG change analyses and artificial intelligence algorithm training or external validation [37,38].

As a first step, the present study aims to describe the electrocardiographic phenotypes of a representative subset of the adult (mainly 18–69 years old) French general population. The analyses have also been stratified by sex and age categories. The ECGs recorded at inclusion in the CONSTANCES cohort have been automatically analysed and adjudicated.

2. Methods

2.1. ECG recording and processing

Our research project was approved by the CONSTANCES scientific committee (project #2022-A-189). All subjects at inclusion in the CONSTANCES cohort were scheduled to have a digital 12-lead, 10-second ECG recording (acquired for the vast majority with a Cardionics ECG recorder). There were no exclusion criteria.

ECG recording was standardized using standard operating procedures. Training and monthly monitoring of nurses in each of the 21 health screening centres was implemented by the national social security fund (Caisse Nationale de l'Assurance Maladie). Epidemiological research assistants make regular site visits to health screening centres to check compliance with these standard operating procedures. They also make regular quality checks of all equipment used in these centres [39].

All ECG files were anonymized and stored in public domain format (either the SCP or the HL7 XML standard) in the CONSTANCES database. For the present study, only ECG waveforms, subject sex and age demographics were considered.

All digital ECG files were subsequently automatically analysed using commercial software that embeds the Glasgow diagnostic algorithm (AMPS LLC, New York, NY) [40,41]. The Glasgow method and the automatic output of the software are described in Text A.1.

2.2. Adjudication of automated measurements and diagnoses

Data obtained from the automated analysis were used to describe the prevalence and distribution of all parameters and for ECG selection for the adjudication process.

2.2.1. Adjudication process

Two senior cardiologists (F.E. and P.M.-B.; cardiologists 1 and 2, respectively) performed the adjudication, which consisted of a batch qualitative and/or quantitative adjudication of each ECG parameter. Selected ECGs were displayed using the CalECG software (version 4.1.0, AMPS LLC, NY) (Fig. 1A). Automated lead-independent time interval measurements (PR, QRS and QT intervals) was based on the 12-lead overlap medians (Fig. 1B and C). Adjudication of automated time interval measurements was performed using prepositioned vertical moving callipers for P onset, QRS onset, QRS offset and QT end. Interval durations were calculated between relevant points.

Adjudication of automated morphological analysis was performed on screen using different ECG display formats (mostly the 6 × 2 or 4 × 3 configurations). Adjudication of rhythm analysis was performed using on-screen 10-second displays of beat-to-beat ECG tracings.

ECG abnormalities with different adjudications by cardiologists 1 and 2 have subsequently been reconciled by consensus. The adjudication of abnormal parameters was intended to estimate the number of false positive abnormality diagnoses. In order to estimate the number of false negative diagnoses, we also adjudicated a randomly selected subset of 1% of ECGs automatically diagnosed as normal.

2.2.2. Adjudicated parameters

All ECGs with the following quantitative criteria were adjudicated:

- Ventricular rate < 45 or > 100 bpm;
- PR interval < 120 or > 220 ms;
- QRS interval < 80 or > 120 ms;
- Bazett's and Fridericia's corrected QT intervals (QTcB and QTcF, respectively) < 350 or > 480 ms;
- QRS axis < –30° or > 110°.

ECGs with abnormal qualitative diagnostics were adjudicated, considering a subset of different statements. When a given abnormality was detected in more than 1% of ECGs, the adjudication was limited to a maximum of 1% of the total number of ECGs (the first 1400 in numerical order).

The following statements were adjudicated:

- Premature beats (A, J or V);
- Supraventricular tachycardia (atrial tachycardia, atrial flutter, atrial fibrillation);
- Ventricular arrhythmias (ventricular tachycardia, accelerated idiopathic ventricular rhythm, ventricular flutter, Torsade de Pointes, ventricular fibrillation);
- Conduction defects (second- and third-degree sinoatrial block, first-, second- and third-degree atrioventricular block, right bundle branch block [RBBB], left bundle branch block [LBBB], non-specific intraventricular conduction defect);
- QRS complex morphology;
- Acute myocardial infarction, chronic myocardial infarction;
- ST and T wave morphology;
- Abnormal repolarization morphology (abnormal ST segment, T or U waves).

2.3. Analyses and statistics

Qualitative parameters are presented as values and percentages. Agreement between over-readers was estimated by Cohen's kappa test. Quantitative data are presented as means ± standard deviations (SDs), after checking for their normal distribution, and as percentiles.

For each parameter, automated detection of abnormal parameters (intervals or diagnosis) were adjudicated to detect 'false positive' findings. We then calculated the number of 'true positive' parameter = automated detection minus false positive identified by adjudication.

In addition, a 1% subset of randomly selected ECG automatically diagnosed as normal was adjudicated. Each ECG showing an abnormal parameter during adjudication was considered as a 'false negative' ECG.

The 'estimated prevalence' for each parameter was calculated as follows: estimated prevalence = (number of true positives + (number of false negatives * 100))/total number of ECGs.

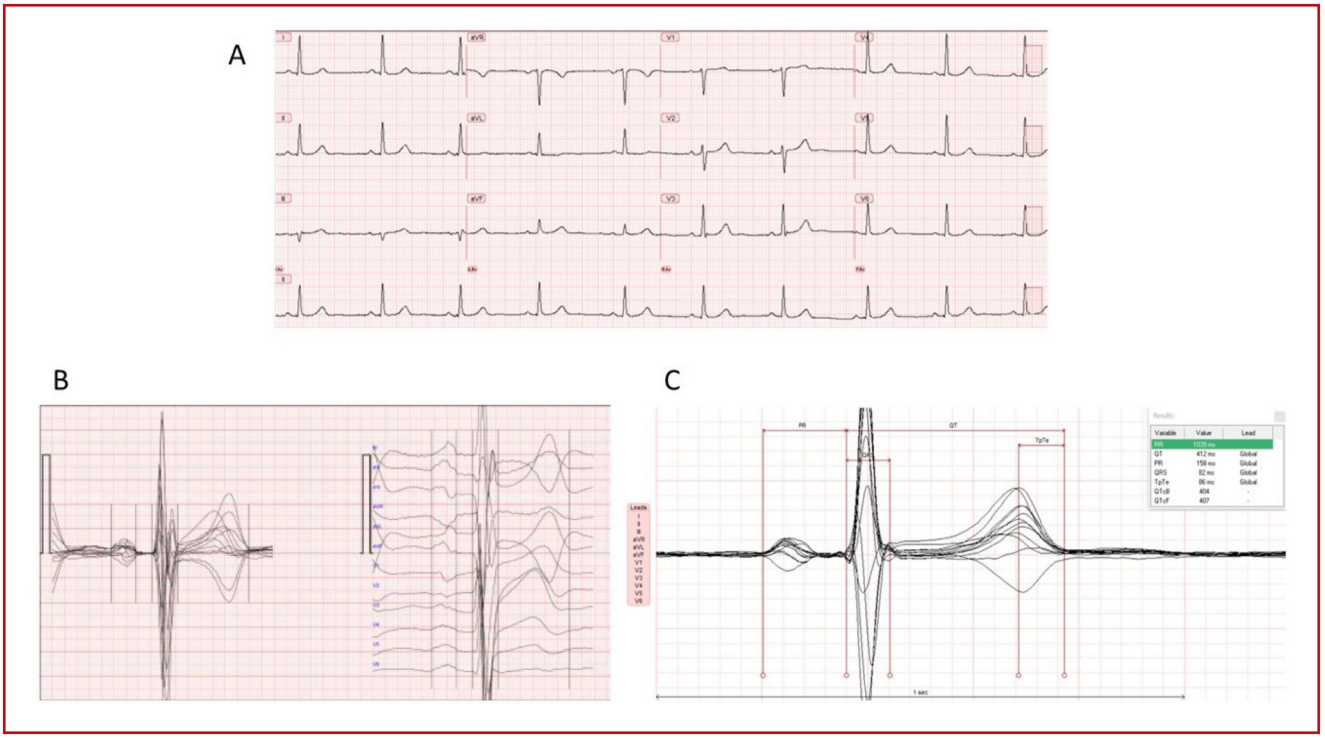


Fig. 1. XML ECG displays. A. ECGs were displayed using the CalECG software (version 4.1.0, AMPS LLC, New York, NY). B. 12-lead overlap medians. C. Adjudication of automated time interval measurements performed by moving prepositioned vertical moving cursors for P onset, QRS onset, QRS offset and QT end. ECG: electrocardiogram; QTcB: Bazett's corrected QT interval; QTcF: Fridericia's corrected QT interval.

Of note, the number of false negatives was multiplied by 100 to account for the 1% subset of normal ECGs that were adjudicated.

Age distribution by sex was compared using a Student's test. The ECG measurements were analysed according to sex and age category (18–29, 30–39, 40–49, 50–59, 60–69 and ≥70 years) using two-way analysis of variance (ANOVA). The effect of the RR interval on PR, QRS and corrected QT (QTc) intervals was assessed by linear regression.

The threshold for statistical significance was set at $P < 0.05$. All statistical analyses were performed using RStudio software (Version 4.1.2, 2009-2022 RStudio©, PBC).

3. Results

3.1. Study population

A total of 143,763 digital ECGs recorded in 143,763 subjects (54% female; mean ± SD age 47.0 ± 13.5 years in females and 46.9 ± 13.5 years in males; $P = 0.44$) were available and analysed. Fig. 2 shows the age and sex distributions of subjects included in the study.

3.2. Quantitative ECG measurements

3.2.1. Automated global measurements

Normality of continuous variables distribution was confirmed in all cases. Table 1 displays the distribution of automatically measured parameters. Fig. 3 shows the distribution of RR intervals and QTcF values.

3.2.2. Adjudication for quantitative ECG measurements

A total of 7952 ECGs (6643 with abnormal values and 1309 with normal values) were adjudicated (Table 2). The main reasons for false automatic measurements are listed in Text A.2.

In order to estimate the proportion of false negative abnormal measurements, 1309 normal ECGs (1% of the 130,888 ECGs automatically classified as 'normal') were adjudicated. The numbers of ECGs with false negative (= undetected) abnormal interval values are indicated in Table 2. Out of the 1309 adjudicated normal ECGs, nine had measurements out of the defined ranges (two with PR < 120 ms, two with QRS < 80 ms and five with QTcB < 350 ms, of which two also had QTcF < 350 ms). As 1% of the normal ECGs were adjudicated, the total number of false negatives was estimated by multiplying the related value by 100. The number of true positives + false negatives × 100 was used to calculate the estimated prevalence of abnormal interval values (Table 2).

Adjudication for intervals (i.e. interval long or short YES/NO) was performed by two over-readers (cardiologists 1 and 2) for QTc values. The inter-rater reliability was intermediate (K coefficient < 0.8) (Table A.1).

Due to the small proportion of measurement errors, exclusion of measurement errors did not change the overall distribution of quantitative parameters (Table A.2).

3.2.3. Clinical correlates: Age, sex and RR effects on adjudicated quantitative ECG parameters

Table 3 shows the distribution of quantitative parameters according to sex and age. Heart rate and interval durations were dependent on both sex and age. The two-way ANOVA was significant for age ($P < 0.0001$ for heart rate and PR, QRS, QTcB and QTcF intervals) and sex ($P < 0.0001$ for heart rate and PR, QRS, QTcB and QTcF intervals). In addition, the interaction (sex*age) was also significant ($P < 0.0001$ for heart rate and PR, QRS, QTcB and QTcF intervals). When compared to males, females had faster heart rates and longer QTc interval durations but shorter PR and QRS interval durations. PR, QRS and QTc interval durations increased with age in both sexes.

PR, QRS, QTcB and QTcF durations were all significantly but weakly correlated to the RR interval (R^2 0.02, 0.02, 0.03 and 2×10^{-5}, respectively). Due to its lowest correlation to RR interval, the Fridericia correction appears to be less biased than that of Bazett.

Fig. 2. Age and sex distribution of subjects included in the study.

Table 1
Automated quantitative ECG measurements (n = 143,763).

Parameter	Mean ± SD (95% CI)	Percentiles (1st; 2nd)	Percentiles (98th; 99th)
Heart rate (bpm)	63 ± 10 (62.9–63.1)	43; 45	88; 93
RR interval (ms)	964 ± 156 (963.2–964.8)	643; 678	1328; 1389
P wave duration (ms)	108 ± 13 (107.9–108.1)	70; 76	132; 136
P frontal axis (degrees)	45 ± 27 (44.9–45.1)	–37; –23	80; 82
PR interval (ms)	156 ± 24 (155.9–156.1)	106; 112	210; 222
QRS duration (ms)	94 ± 11 (93.9–94.1)	74; 76	122; 134
QRS frontal axis (degrees)	38 ± 33 (37.8–38.2)	–51; –38	90; 96
ST segment duration (ms)	113 ± 28 (112.9–113.1)	32; 46	164; 170
ST frontal axis (degrees)	6 ± 95 (5.5–6.5)	–174; –167	170; 176
T-wave duration (ms)	201 ± 26 (200.9–201.1)	150; 156	262; 278
T-wave frontal axis (degrees)	37 ± 22 (36.9–37.1)	–16; –6	75; 85
QT interval duration (ms)	408 ± 29 (407.9–408.1)	344; 352	470; 482
QTc Hodges (ms)	414 ± 19 (413.9–414.1)	374; 379	457; 465
QTcB (ms)	416 ± 22 (415.9–416.1)	365; 371	461; 468
QTcF (ms)	413 ± 19 (412.9–413.1)	372; 377	453; 461
QTc Framingham (ms)	412 ± 19 (411.9–412.1)	368; 374	452; 460

CI: confidence interval; ECG: electrocardiogram; QTc: corrected QT; QTcB: Bazett's corrected QT interval; QTcF: Fridericia's corrected QT interval; SD: standard deviation.

3.3. ECG diagnosis statements

3.3.1. Automated ECG diagnoses

Table 4 shows the prevalences of the different rhythms and abnormal morphological parameters automatically diagnosed by the algorithm. More diagnostic categories are provided in Tables A.3 and A.4.

3.3.2. Adjudication for automated ECG diagnosis statements

A total of 4319 ECGs (3010 with abnormal diagnoses and 1309 considered as normal by the algorithm) were adjudicated for qualitative diagnosis. Table A.5 shows agreement between cardiologists 1 and 2.

Table 5 shows the adjudicated and estimated prevalences of ECG diagnoses. Among the 1309 normal ECGs that were adjudicated for any abnormality, at least one qualitative abnormality on 121 (9.2%) ECGs was found. Automated diagnosis errors included: 39 ST-T non-specific abnormalities (32.2% of abnormalities), 33 axis deviations (27.3%), 23 premature beats (19.0%), 11 first-degree atrioventricular block (> 200 ms) (9.1%), nine incomplete RBBB (7.4%) and 11 other abnormalities (9.1%) including three ectopic atrial rhythms, three atrial hypertrophy, two low QRS voltage, one early repolarization pattern, one R wave increase on lead V1 and one lead error.

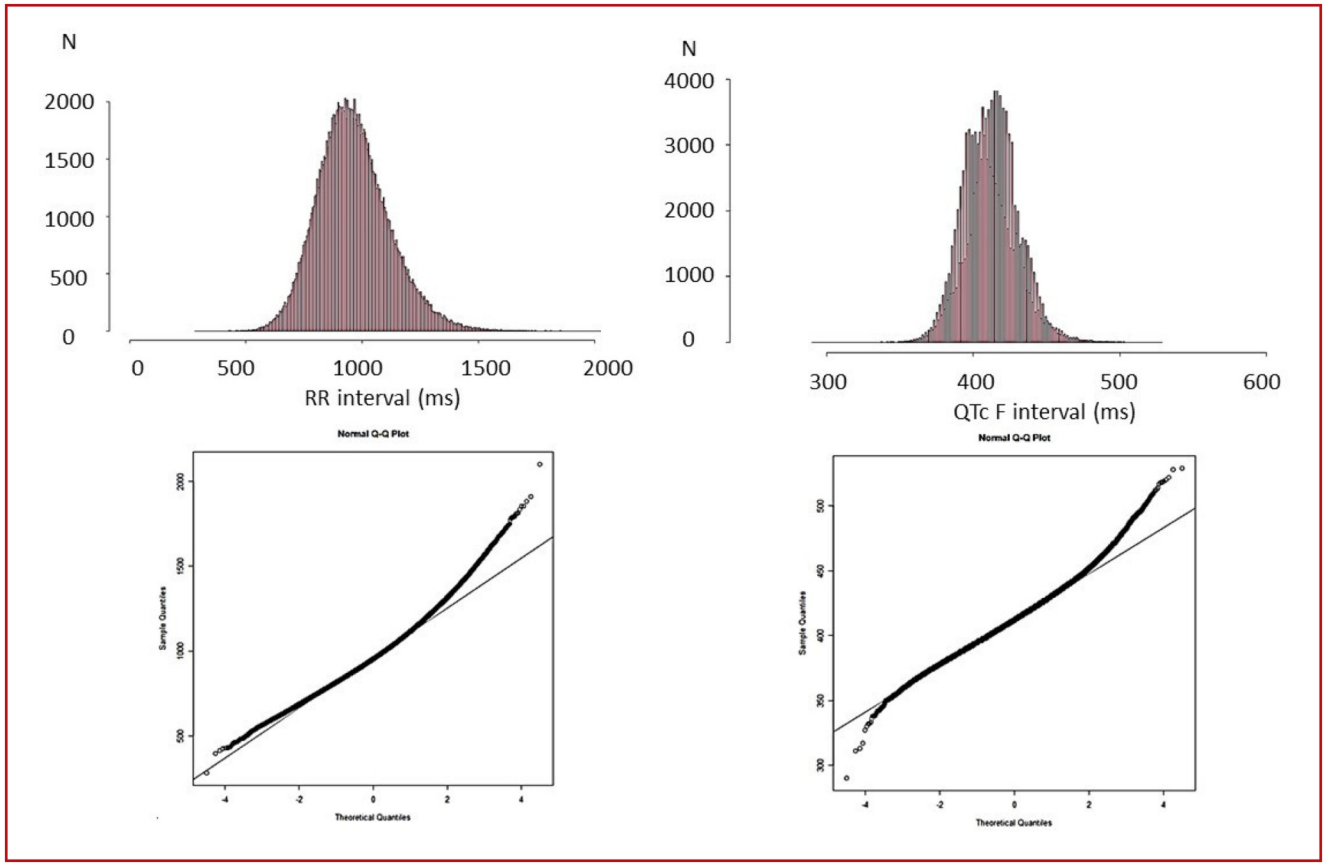


Fig. 3. Distribution of RR intervals and QTcF values in the entire cohort. QTcF: Fridericia's corrected QT interval.

Table 2
Adjudicated and estimated prevalences of out-of-range quantitative ECG measurements.

Interval (ms)	Automated measurement (out of 143,763)	False positive (rejected by Adjudication)	True positive (confirmed after adjudication)	False negative (out of 1309)	Estimated prevalence[a] (%)
PR < 120	641	179	462	2	0.46 (0.43–0.50)
PR > 220	1570	126	1444	0	1.00 (0.95–1.06)
QRS < 80	224	17	207	2	0.28 (0.26–0.31)
QRS > 120	2956	592	2364	0	1.64 (1.58–1.71)
QTcB < 350	338	14	324	5	0.57 (0.53–0.61)
QTcB > 480	530	241	289	0	0.20 (0.18–0.22)
QTcF < 350	143	29	114	2	0.22 (0.19–0.24)
QTcF > 480	241	107	134	0	0.09 (0.08–0.11)

Data are expressed as numbers or % (95% confidence interval). ECG: electrocardiogram; QTcB: Bazett's corrected QT interval; QTcF: Fridericia's corrected QT interval.
[a] Estimated prevalence = (true positive + (false negative × 100))/total number of ECGs.

4. Discussion

4.1. Main findings

We analysed the ECGs of 143,763 subjects at inclusion in one of the largest longitudinal ECG cohorts in the general population. A total of 9653 ECGs with abnormal automated computerized analysis of qualitative diagnoses and quantitative parameters, as well as a subset of 1% of ECGs that had automatically been classified as 'normal' ($n = 1309$), were adjudicated (in total, there were > 10,000 adjudicated ECGs). We could hence describe the electrocardiographic phenotypes of a representative subset of the adult (mainly 18–69 years old) French general population.

4.2. Epidemiological ECG cohorts

ECG phenotypes have been described from cohorts of increasing size over time (from $n = 503$ [14] to more than 3.5 million subjects [33,34]). Among the seven studies that have included more than 100,000 subjects, one study included healthy subjects [23], three included primary care patients [25–27], one included unselected subjects [28] and two retrieved ECGs from healthcare or health insurance association databases [33,34].

While not the largest in terms of the number of subjects, the CONSTANCES cohort has included more than 200,000 subjects, 143,763 of whom had a baseline digital ECG (the ECG was not digitally recorded in the remaining subjects included in the cohort). The novelty of the CONS-

Table 3
Distribution of quantitative parameters according to sex and age.

Age group (years)	Heart rate (bpm)	PR interval (ms)	QRS interval (ms)	QTcB (ms)	QTcF (ms)
Males					
18–29 (n = 8566)	61 ± 11 (41–92)	150 ± 21 (104–210)	98 ± 10 (78–126)	399 ± 20 (350–447)	399 ± 17 (360–442)
30–39 (n = 13,782)	60 ± 10 (40 -89)	156 ± 22 (110–216)	98 ± 10 (78–126)	402 ± 20 (355–450)	403 ± 17 (366–445)
40–49 (n = 15,614)	61 ± 10 (41–90)	160 ± 22 (112–222)	98 ± 10 (78–126)	407 ± 20 (361–456)	407 ± 17 (370–451)
50–59 (n = 13,978)	62 ± 11 (42–94)	164 ± 23 (114–230)	98 ± 11 (78–144)	413 ± 20 (367–465)	411 ± 17 (375–459)
60–69 (n = 14,305)	63 ± 11 (43–96)	170 ± 26 (114–248)	99 ± 14 (76–154)	418 ± 21 (371–479)	416 ± 19 (377–469)
70–73 (n = 520)	63 ± 12 (43–98)	174 ± 29 (109–263)	101 ± 15 (77–159)	421 ± 22 (376–494)	418 ± 19 (379–472)
Females					
18–29 (n = 9755)	66 ± 11 (45–96)	143 ± 20 (100–214	89 ± 8 (72–120	416 ± 19 (370–470	410 ± 17 (373–462
30–39 (n = 16,010)	64 ± 10 (44–91)	147 ± 21 (102–204	90 ± 9 (72–114	419 ± 19 (372–465	414 ± 17 (376–457
40–49 (n = 17,918)	65 ± 10 (45–93)	149 ± 22 (104–208	90 ± 9 (72–116	421 ± 19 (375–467	416 ± 17 (379–460
50–59 (n = 16,080)	64 ± 10 (45–92)	156 ± 22 (108–216	91 ± 10 (72–124	424 ± 19 (378–471	419 ± 17 (382–463
60–69 (n = 16,543)	65 ± 10 (46–94)	160 ± 23 (110–223	91 ± 11 (72–136	427 ± 19 (382–480	421 ± 18 (384–471
70–73 (n = 626)	66 ± 10 (46–92)	163 ± 25 (113–236	92 ± 13 (75–150	428 ± 20 (388–488	422 ± 18 (383–473

Data are expressed as mean ± standard deviation (1^{st}–99^{th} percentiles). QTcB: Bazett's corrected QT interval; QTcF: Fridericia's corrected QT interval.

Table 4
Prevalence of automatic diagnoses among the analysed ECGs.

	All analysed ECGs (n = 143,763)
Rhythm diagnosis statements	
Sinus rhythm	140,598 (97.8) [97.7–97.9]
Ectopic atrial rhythm	2308 (1.6) [1.54–1.67]
Supraventricular tachycardia	437 (0.3) [0.28–0.33]
Junctional rhythm	189 (0.13) [0.11–0.15]
Other arrhythmias	231 (0.16) [0.14–0.18]
Morphological parameter statements	
Premature beats	3017 (2.1) [2.0–2.2]
Abnormal P wave morphology	1001 (0.7) [0.65–0.67]
Abnormal atrioventricular conduction	6954 (4.8) [4.7–4.9]
Wolff-Parkinson-White pattern	92 (< 0.1) [0.05–0.08]
Ventricular conduction defect	11,403 (7.9) [7.8–8.1]
Abnormal QRS axis	8511 (5.9) [5.8–6.0]
Abnormal QRS voltage	2399 (1.7) [1.6–1.7]
J wave pattern	173 (0.12) [0.10–0.14]
Abnormal QT interval	1389 (0.97) [0.92–1.02]
Acute myocardial infarction	1397 (0.97) [0.92–1.02]
Chronic myocardial infarction	11,235 (7.8) [7.7–7.9]
Abnormal ST-T	33,082 (23.0) [22.8–23.2]
Pacemaker	1299 (0.9) [0.85–0.95]
Diagnosis conclusion	
Normal ECG	78,680 (54.7) [54.5–55.0]
Abnormal ECG	12,872 (9.0) [8.8–9.1]
Borderline ECG	45,785 (31.8) [31.6–32.1]
Normal ECG based on available leads	632 (0.44) [0.41–0.47]
Normal ECG except for heart rate	5766 (4.0) [3.9–4.1]
Missing data	28 (0.02) [0.01–0.03]

Data are expressed as number (%) [95% confidence interval]. ECG: electrocardiogram.

TANCES cohort lies in its specific epidemiological design. CONSTANCES is a population-based cohort designed to be representative of the French adult population [36]. Accordingly, the ECGs were not recorded for a specific reason or in a specific setting and the study population was not limited to healthy subjects. Of note, studies including ECGs from more than 3 million subjects have not provided quantitative data measurements [33,34]. Hence, our study is the largest ECG study with both qualitative statements and parameter measurements in a representative subset of the population. The adjudication of abnormal automated results is also unique to our study.

The CONSTANCES cohort will allow the scientific community to cross-analyse ECG data with health data (from questionnaires and physiological measurements) and administrative data and professional and residential environment data. In addition, subjects included in the CONSTANCES cohort will have serial ECG recordings during prolonged follow-up. Included in such a research and public health infrastructure, baseline ECG data and ECG changes will be useful to select subgroups of subjects according to diseases and/or exposure to risk factors, but also to potentially predict the future occurrence of events.

4.3. Automated measurements and adjudication

Massive longitudinal digital ECG databases cannot be manually measured and visually interpreted. Digital recording and storage of ECGs has opened the opportunity for automated ECG measurements and interpretation [6–8]. For the present study, we used the Glasgow diagnostic algorithm, which was developed by Macfarlane et al. [40]. This programme has been validated in academic settings, is not dependent on a specific ECG machine vendor and can be run on XML ECG format files [42]. In addition, we need to keep in mind that different algorithms for automated ECG analyses and different ECG lead settings may provide different measurements and interpretation even in the artificial intelligence era [8,42–46].

With these limitations in mind, we decided to adjudicate ECGs with out-of-range values and abnormal interpretations and to look for false negative diagnoses in a 1% subset of ECGs that were automatically classified as normal. Despite using a user-friendly batch process, the adjudication of > 10,000 ECGs by two senior cardiologists (F.E. and P.M.-B.) represented a significant burden. It was, however, important for the scientific community interested in using the CONSTANCES ECG database to be aware of the potential limitations of the automated measurement and interpretation of ECGs. On the other hand, the false positive abnormal ECG measurements or diagnostic errors were, in most cases, below 0.1%. False negatives were high (1.8%) for 'other arrhythmias' mainly because of premature beats and sinus arrhythmias considered as normal by the algorithm but not by the cardiologists. For other measurements and diagnoses, false negatives were quite small and, in most cases, due to values close to cut-off definitions rather than to significant errors. Overall, although selection of a subpopulation based on ECG parameters may warrant ECG adjudication, the automated measurement and interpretation provided by the Glasgow algorithm proved highly efficient for epidemiological evaluation.

4.4. Quantitative parameter measurements

Mean values of interval measurements (heart rate and PR, QRS and QTc intervals) were comparable to those described in healthy subjects

Table 5
Adjudicated and estimated prevalences of ECG diagnoses.

Parameter	Automated diagnosis (out of 143,763)	False positives (rejected by adjudication)	True positives (confirmed after adjudication)	False negatives (out of 1309)	Estimated prevalence (%)
Sinus tachycardia	462	10	452	0	0.31 (0.29–0.34)
Marked sinus bradycardia	102	0	102	0	0.07 (0.06–0.08)
Atrial fibrillation	352	136	216	0	0.15 (0.13–0.17)
Atrial flutter	83	67	16	0	0.01 (0.01–0.02)
Other arrhythmias	204	191	13	26	1.82 (1.75–1.89)
Sinoatrial and atrioventricular blocks	17	0	17	11	0.78 (0.73–0.82)
Wolff-Parkinson-White pattern	92	22	70	0	0.05 (0.04–0.06)
RBBB	1304	144	1160	0	0.81 (0.76–0.85)
LBBB	394	35	359	0	0.25 (0.22–0.28)

Data are expressed as number (%) or % (95% confidence interval). ECG: electrocardiogram; LBBB: left bundle branch block; RBBB: right bundle branch block.

[15,24], but also in large unselected population studies [20,28] as well as in primary care [25].

Sex- and age-related differences in parameter measurements (faster heart rates and longer QTc interval durations but shorter PR and QRS interval durations in females and PR, QRS and QTc interval duration increases with age in both sexes) are consistent with those previously reported in the literature [15,24,25]. The reasons for such differences are complex. Steroid hormones play a critical role in sexual differences in ventricular repolarization duration [47]. It is also known that the level and impact of steroid hormones change with aging. In addition, both age and sex influence the autonomic regulation of heart rate and thereby may impact heart rate and ECG interval durations [48]. Both steroid hormones and autonomic influences also interact with structural aging of the heart.

Estimation of abnormal parameter durations after adjudication provides further information (Table 2). PR interval is generally considered as abnormal when it is < 120 ms in adults. This was observed in 0.5% of subjects in our population, after exclusion of Wolff-Parkinson-White pattern. Similarly, a short QRS interval duration was found in 0.3% of subjects. We will discuss conduction defects in the next paragraph.

Regarding QT/QTc interval duration, the QTcF had the lowest correlation with the RR interval. Accordingly, and as described by others [49,50], the Fridericia correction formula appears less biased than the Bazett correction formula in general population settings. QTcF was associated with a lower prevalence of prolonged (> 480 ms) or short (< 350 ms) QTc intervals (Table 2). However, we previously showed that the QT rate dependence was lower in patients with short QT syndrome when compared with control subjects [51]. Accordingly, the Fridericia correction formula would overestimate the QTc duration and thus underestimate the prevalence of short QTc interval. The QTcF interval was < 350 ms in 0.2% of the subjects in our population. This prevalence is slightly below the 0.3% of QTcF < 340 ms in the Finnish cohort but within the confidence interval [52]. Of note, short QTc values were not associated with increased mortality in the Finnish cohort [52].

4.5. ECG diagnosis statements

The prevalence of adjudicated ECG diagnoses reported in Table 5 is very similar to those described in the healthy subjects included in the Dutch LifeLines cohort study [23]. In our study, the prevalence of atrial fibrillation was < 0.2%. The recording of a single 10-second ECG most probably misses paroxysmal atrial fibrillation episodes and hence underestimates the true prevalence of atrial fibrillation. However, such a prevalence is in accordance with data from the Global Burden of Disease Study [53] when considering the relatively young age (mean 47 ± 14 years) of subjects included in the CONSTANCES cohort. Conversely, data from half a million adults included in the United Kingdom biobank indicate a prevalence of atrial fibrillation of 1.54% (95% confidence interval 1.50–1.57) [54]. This higher prevalence could be the consequence of a slightly older age (median age 58 years), but even in the subgroup of subjects aged < 55 years, the prevalence of atrial fibrillation was 1.4% in males and 0.6% in females [54]. Even taking into account the underestimation of atrial fibrillation prevalence using only one 10-second ECG recording, these results suggest different atrial fibrillation prevalences in France and the United Kingdom.

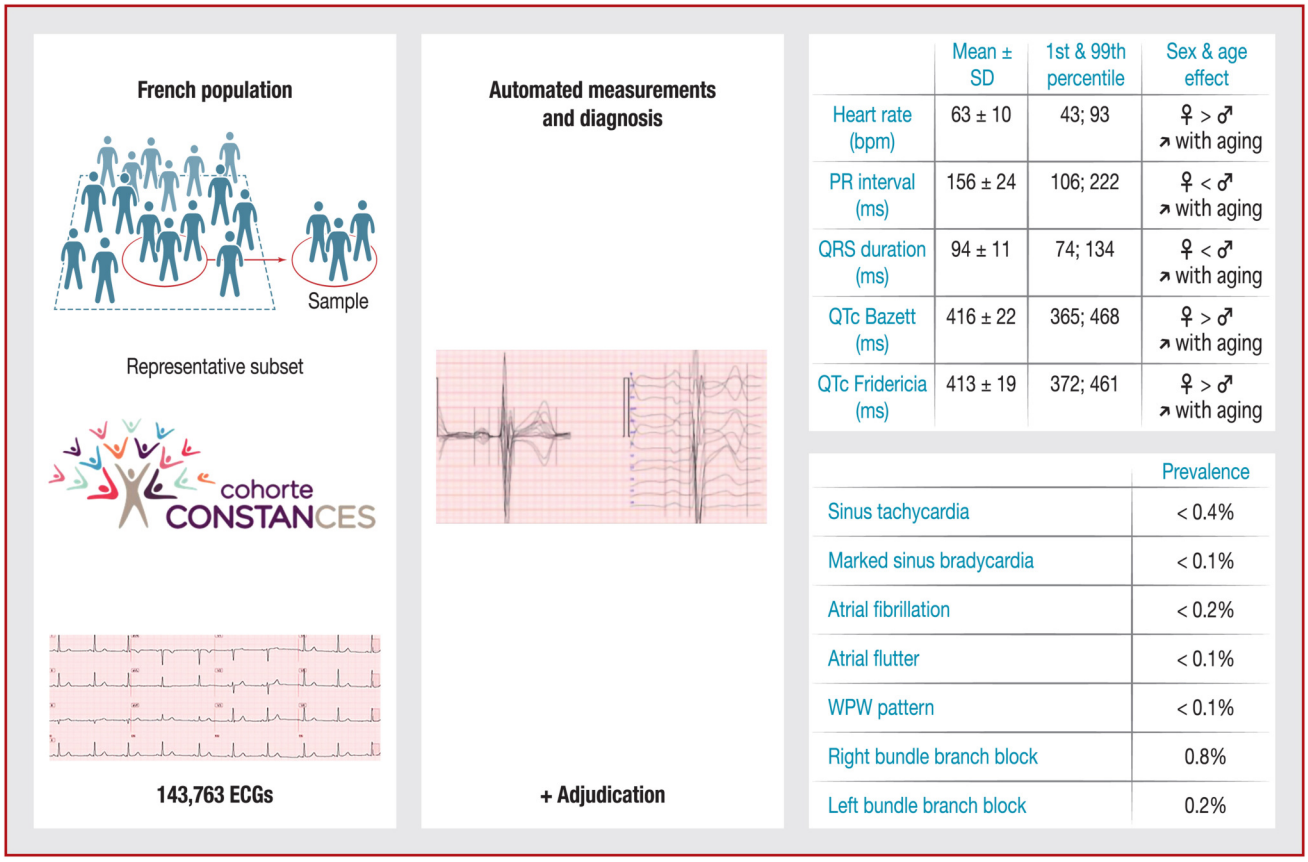
The prevalences of RBBB (0.8%) and LBBB (0.2%) in the CONSTANCES cohort is also in line with data from the LifeLines cohort study [23]. Other studies describe higher prevalences of conduction defects. The prevalence of RBBB was 3.2% in patients from 29 urban primary healthcare centres in the Barcelona city area [55]. Data from the Mass General Brigham multi-institutional healthcare system describe a prevalence of infra-nodal conduction defect (including bundle branch blocks but also fascicular blocks and non-specific conduction defects) of 15% [56]. Although these two studies [55,56] included slightly older subjects than the CONSTANCES cohort, the differences in conduction defect prevalence suggest an overrepresentation of such abnormalities in primary care setting when compared to general population settings.

4.6. Limitations

While manual adjudication of the abnormal parameters is a strength of our study, the burden of adjudication prevented the adjudication of all abnormal parameters. In addition, adjudicating only 1% of ECGs automatically labelled as 'normal' could have underestimated false negatives for rare abnormalities and hence their estimated prevalences.

5. Conclusions

We described ECG phenotypes in one of the largest representative cohorts of the general population mainly aged 18–69 years. We showed that ECG quantitative parameters and diagnosis prevalences were comparable to those observed in other apparently healthy cohorts. Conversely, ECG abnormalities were less prevalent when compared to ECG data collected from primary care cohorts. This ECG phenotype characterization will open up possibilities for cross-analyses with medical, sociological and environmental data gathered in the CONSTANCES cohort. In addition, the representativeness of this ECG cohort will make it precious for learning, but also external validation for artificial intelligence-based algorithms [57,58]. Finally, this first set of ECGs–recorded at inclusion–will serve as a reference to evaluate serial ECG changes during extended follow-up. These future steps may lay the groundwork for future personalized predictive applications (Central illustration).



	Mean ± SD	1st & 99th percentile	Sex & age effect
Heart rate (bpm)	63 ± 10	43; 93	♀ > ♂ ↗ with aging
PR interval (ms)	156 ± 24	106; 222	♀ < ♂ ↗ with aging
QRS duration (ms)	94 ± 11	74; 134	♀ < ♂ ↗ with aging
QTc Bazett (ms)	416 ± 22	365; 468	♀ > ♂ ↗ with aging
QTc Fridericia (ms)	413 ± 19	372; 461	♀ > ♂ ↗ with aging

	Prevalence
Sinus tachycardia	< 0.4%
Marked sinus bradycardia	< 0.1%
Atrial fibrillation	< 0.2%
Atrial flutter	< 0.1%
WPW pattern	< 0.1%
Right bundle branch block	0.8%
Left bundle branch block	0.2%

Central illustration. ECG phenotypes in the CONSTANCES cohort (a representative subset of the French general population). ♀ : female; ♂ : male; ECG: electrocardiogram; QTcB: Bazett's corrected QT interval; QTcF: Fridericia's corrected QT interval; SD: standard deviation.

Sources of funding

None.

Acknowledgements

None.

Disclosure of interest

The authors declare that they have no competing interest.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.acvd.2025.12.013.

References

[1] GBD 2021 Causes of Death Collaborators. Global burden of 288 causes of death and life expectancy decomposition in 204 countries and territories and 811 subnational locations, 1990-2021: a systematic analysis for the Global Burden of Disease Study 2021. Lancet 2024;403(10440):2100–32, http://dx.doi.org/10.1016/S0140-6736(24)00367-2.
[2] Global NCD Compact 2020–2030. (http://www.who.int).
[3] Kligfield P, Gettes LS, Bailey JJ, et al. Recommendations for the Standardization and Interpretation of the Electrocardiogram. Part I: The Electrocardiogram and Its Technology. A Scientific Statement From the American Heart Association Electrocardiography and Arrhythmias Committee, Council on Clinical Cardiology; the American College of Cardiology Foundation; and the Heart Rhythm Society. Endorsed by the International Society for Computerized Electrocardiology. Circulation 2007;115:1306–24.
[4] Curry SJ, Krist AH, Owens DK, et al. Screening for cardiovascular disease risk with electrocardiography: US Preventive Services Task Force Recommendation Statement. US Preventive Service Task Force. JAMA 2018;319:2308–14.
[5] Davidson KW, Barry MJ, Mangione CM, et al. Screening for atrial fibrillation: US preventive services task force recommendation statement. US Preventive Services Task Force. JAMA 2022;327:360–7.
[6] Macfarlane PW, et al. Minnesota coding and the prevalence of ECG abnormalities. Heart 2000;84(6):582–4, http://dx.doi.org/10.1136/heart.84.6.582.
[7] Ball RL, Feiveson AH, Schlegel TT, Starc V, Dabney AR. Predicting "Heart Age" using electrocardiography. J Personal Med 2014;4:65–78.
[8] Kligfield P, Badilini F, Denjoy I, et al. Comparison of automated interval measurements by widely used algorithms in digital electrocardiographs. Am Heart J 2018;200:1–10.
[9] Macfarlane PW, Latif S. Automated serial ECG comparison based on the Minnesota code. J Electrocardiol 1996;29(Suppl.):29–34.
[10] Prifti E, Fall A, Davogustto G, et al. Deep learning analysis of electrocardiogram for risk prediction of drug-induced arrhythmias and diagnosis of long QT syndrome. Eur Heart J 2021;42:3948–61.
[11] Ulloa-Cerna AE, Jing L, Pfeifer JM, et al. rECHOmmend: An ECG-based machine learning approach for identifying patients at increased risk of undiagnosed structural heart disease detectable by echocardiography. Circulation 2022;146:36–47.
[12] Khurshid S, Friedman S, Reeder C, et al. ECG-based deep learning and clinical risk factors to predict atrial fibrillation. Circulation 2022;145:122–33.
[13] Sangha V, Nargesi AA, Dhingra LS, et al. Detection of left ventricular systolic dysfunction from electrocardiographic images. Circulation 2023;148:765–77.
[14] Chen CY, Chiang BN, Macfarlane PW. Normal limits of the electrocardiogram in a Chinese population. J Electrocardiol 1989;22:1–15.
[15] Wu J, Kors JA, Rijnbeek PR, van Herpen G, Lu Z, Xu C. Normal limits of the electrocardiogram in Chinese subjects. Int J Cardiol 2003;87:37–51.
[16] Katibi I, Clark EN, Devine B, Lloyd SM, Macfarlane PW. Normal limits of the electrocardiogram in Nigerians. J Electrocardiol 2013;46:289–95.
[17] Rautaharju PM, Zhang ZM, Gregg RE, et al. Normal standards for computer-ECG programs for prognostically and diagnostically important ECG variables derived from a large ethnically diverse female cohort: TheWomen's Health Initiative (WHI). J. Electrocardiol 2013;46:707–16.
[18] Dewhurst MJ, Di Marco LY, Dewhurst F, et al. Electrocardiographic reference values for a population of older adults in sub-Saharan Africa. Ann Noninvasive Electrocardiol 2014;19:34–42.

[19] Macfarlane PW, Katibi IA, Hamde ST, et al. Racial differences in the ECG–selected aspects. J Electrocardiol 2014;47:809–14.
[20] Rijnbeek PR, van Herpen G, Bots ML, et al. Normal values of the electrocardiogram for ages 16-90 years. J Electrocardiol 2014;47:914–21.
[21] Macfarlane PW, Lloyd SM, Singh D, et al. Normal limits of the electrocardiogram in Indians. J Electrocardiol 2015;48:652–68.
[22] Bratincsák A, Kimata C, Limm-Chan BN, Vincent KP, Williams MR, Perry JC. Electrocardiogram standards for children and young adults using Z-scores. Circ Arrhythmia Electrophysiol 2020;13:e008253.
[23] van der Ende Yldau M, et al. Population-based values and abnormalities of the electrocardiogram in the general Dutch population: the LifeLines Cohort Study. Clin Cardiol 2017;40:865–72.
[24] Mason JW, Ramseth DJ, Chanter DO, Moon TE, Goodman DB, Mendzelevski B. Electrocardiographic reference ranges derived from 79,743 ambulatory subjects. J Electrocardiol 2007;40:228–34.
[25] Palhares DMF, Marcolino MS, Santos TMM, et al. Normal limits of the electrocardiogram derived from a large database of Brazilian primary care patients. BMC Cardiovasc Disord 2017;17:152.
[26] Silva M, Palhares D, Ribeiro L, et al. Prevalence of major and minor electrocardiographic abnormalities in one million primary care Latinos. J Electrocardiol 2020;64:36–41.
[27] Haimovich JS, Di Achille P, Nauffal V, et al. Frequency of electrocardiogram-defined cardiac conduction disorders in a multi-institutional primary care cohort. JACC Adv 2024;3(7):101004.
[28] Giovanardi P, Vernia C, Tincani E, Giberti C, Silipo F, Fabbo A. Combined effects of age and comorbidities on electrocardiographic parameters in a large non-selected population. J Clin Med 2022;11:3737.
[29] De Bacquer D, De Backer G, Kornitzer M. Prevalences of ECG findings in large population based samples of men and women. Heart 2000;84:625–33.
[30] Rautaharju PM, Zhou SH, Wong S, Calhoun HP, Berenson GS, Prineas R, et al. Sex differences in the evolution of the electrocardiographic QT interval with age. Can J Cardiol 1992;8:690–5.
[31] Rijnbeek PR, Witsenburg M, Schrama E, Hess J, Kors JA. New normal limits for the paediatric electrocardiogram. Eur Heart J 2001;22:702–11.
[32] Marek J, Bufalino V, Davis J, Marek K, Garni A, Stephan W, et al. Feasibility and findings of large-scale electrocardiographic screening in young adults: data from 32,561 subjects. Heart Rhythm 2011;8:1555–9.
[33] Bhatia RS, Bouck Z, Ivers NM, et al. Electrocardiograms in low-risk patients undergoing an annual health examination. JAMA Intern Med 2017;177(9):1326–33, http://dx.doi.org/10.1001/jamainternmed. 2017.2649.
[34] Yagi R, Mori Y, Goto S, Iwami T, Inoue K. Routine electrocardiogram screening and cardiovascular disease events in adults. JAMA Intern Med 2024;184(9):1035–44.
[35] Zins M, Bonenfant S, Carton M, Coeuret-Pellicer M, Guéguen A, Gourmelen J, et al. The CONSTANCES cohort: an open epidemiological laboratory. BMC Public Health 2010;10:479.
[36] Zins M, Goldberg M, CONSTANCES team. The French CONSTANCES population-based cohort: design, inclusion and follow-up. Eur J Epidemiol 2015;30:1317–28.
[37] Desroche LM, Trimaille A, Garban T, Bouleti C. Artificial intelligence in cardiovascular medicine: an exoskeleton for perception, reasoning and action. Arch Cardiovasc Dis 2025;118(11):613–20.
[38] Wan S, Wan F, Dai XJ. Machine learning approaches for cardiovascular disease prediction: a review. Arch Cardiovasc Dis 2025 [S1875-2136(25)00320-1].
[39] Ruiz F, Goldberg M, Lemonnier S, et al. High quality standards for a large-scale prospective population-based observational cohort: Constances. BMC Public Health 2016;16:877.
[40] Macfarlane PW, Devine B, Clark E. The University of Glasgow (Uni-G) ECG analysis program. Comput Cardiol 2005;32:451–4.
[41] https://www.bplmedicaltechnologies.com/blog/glasgow-ecg-interpretation-algorithm-24569/.
[42] Sarapa N, Mortara JL, Brown BD, Isola L, Badilini F. Quantitative performance of E-Scribe warehouse in detecting quality issues with digital annotated ECG data from healthy subjects. J Clin Pharmacol 2008;48(5):538–46.
[43] Hannun AY, Rajpurkar P, Haghpanahi M, et al. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. Nat Med 2019;(1):65–9.
[44] Giudicessi JR, Schram M, Bos JM, et al. Artificial intelligence-enabled assessment of the heart rate corrected QT interval using a mobile electrocardiogram device. Circulation 2021;143(13):1274–86.
[45] Hughes JW, Olgin JE, Avram R, et al. Performance of a convolutional neural network and explainability technique for 12-lead electrocardiogram interpretation. JAMA Cardiol 2021;6(11):1285–95.
[46] Decroocq V, Decroocq M, Riolet C, Dievart F, Pecriaux V, Layec J, et al. Single-lead electrocardiograms and artificial intelligence for managing cardiac rhythm irregularities in general practice: a French general practitioners' survey. Arch Cardiovasc Dis 2025;118(11):582–7.
[47] Salem JE, Alexandre J, Bachelot A, Funck-Brentano C. Influence of steroid hormones on ventricular repolarization. Pharmacol Ther 2016;167:38–47.
[48] Abhishekh HA, Nisarga P, Kisan R, Meghana A, Chandran S, Trichur R, et al. Influence of age and gender on autonomic regulation of heart. J Clin Monitor Comput 2013;27:259–64.
[49] Vandenberk B, Vandael E, Robyns T, et al., Which QT. Correction formulae to use for QT monitoring? Am Heart Assoc 2016;5(6):e003264.
[50] Extramiana F, Maison-Blanche P, Cabanis MJ, Ortemann-Renon C, Beaufils P, Leenhardt A. Clinical assessment of drug-induced QT prolongation in association with heart rate changes. Clin Pharmacol Ther 2005;77(4):247–58.
[51] Extramiana F, Maury P, Maison-Blanche P, Duparc A, Delay M, Leenhardt A. Electrocardiographic biomarkers of ventricular repolarisation in a single family of short QT syndrome and the role of the Bazett correction formula. Am J Cardiol 2008;101:855–60.
[52] Anttonen O, Junttila MJ, Rissanen H, Reunanen A, Viitasalo M, Huikuri HV. Prevalence and prognostic significance of short QT interval in a middle-aged Finnish population. Circulation 2007;116:714–20.
[53] Cheng S, He JZ, Han Y, Han S, Li P, Liao H, et al. Global burden of atrial fibrillation/atrial flutter and its attributable risk factors from 1990 to 2021. Europace 2024;26:euae195.
[54] Khurshid S, Choi SH, Weng LC, et al. Frequency of cardiac rhythm abnormalities in a half million adults. Circ Arrhythm Electrophysiol 2018;11:e006273.
[55] Alventosa-Zaidin M, Guix Font L, Benitez Camps M, et al. Right bundle branch block: prevalence, incidence, and cardiovascular morbidity and mortality in the general population. Eur J Gen Pract 2019;25:109–15.
[56] Haimovich JS, Di Achille P, Nauffal V, et al. Frequency of electrocardiogram-defined cardiac conduction disorders in a multi-institutional primary care cohort. JACC Adv 2024;3:101004.
[57] Villaceque M, Garban T, Lafitte S. Why has a learned society of cardiology created an expert group on artificial intelligence? Arch Cardiovasc Dis 2025;118(11):571–2.
[58] Trimaille A, Lafitte S, Begue F, Beuque G, Weizman O, Lucain P, et al. Unveiling facilitators and barriers to artificial intelligence implementation in cardiac healthcare: rationale and design of the INSIGHT-AI France study, from the Artificial Intelligence Working Group and the National College of Cardiologists in Training of the French Society of Cardiology. Arch Cardiovasc Dis 2025;119(2):132–6.