



HAL
open science

Efficiency of Neonatal Screening for Congenital Adrenal Hyperplasia Due to 21-Hydroxylase Deficiency in Children Born in Mainland France Between 1996 and 2003.

Bénédicte Coulm, Joel Coste, Véronique Tardy, Emmanuel Ecosse, Michel Roussey, Yves Morel, Jean-Claude Carel

► To cite this version:

Bénédicte Coulm, Joel Coste, Véronique Tardy, Emmanuel Ecosse, Michel Roussey, et al.. Efficiency of Neonatal Screening for Congenital Adrenal Hyperplasia Due to 21-Hydroxylase Deficiency in Children Born in Mainland France Between 1996 and 2003.. *Medicine Archives of Pediatrics & Adolescent - JAMA Pediatrics* , 2012, 166 (2), pp.113-120. 10.1001/archpediatrics.2011.774 . inserm-00669946

HAL Id: inserm-00669946

<https://inserm.hal.science/inserm-00669946v1>

Submitted on 14 Feb 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **Title: Efficiency of neonatal screening for congenital adrenal**
2 **hyperplasia due to 21-hydroxylase deficiency in children born in mainland**
3 **France between 1996 and 2003**

4 **Short title:** Screening for congenital adrenal hyperplasia

5 **Authors:** Bénédicte Coulm¹, midwife, Joel Coste¹, MD, Véronique Tardy², MD,
6 Emmanuel Ecosse¹, statistician, Michel Roussey³, MD, Yves Morel², MD, Jean-Claude
7 Carel⁴, MD, on behalf of the DHCSF study group⁴.

8 1 Department of Biostatistics, Groupe hospitalier Cochin - Saint Vincent de Paul and University Paris 5 René
9 Descartes, 75014 Paris, France.

10 2 Department of Biochemistry and Molecular Biology, Molecular Endocrinology and Rare Diseases, CBPE,
11 Groupement hospitalier Lyon-Est, Hospices Civils de Lyon, 59 boulevard Pinel 69677 Bron, France

12 3 Association Française pour le Dépistage et la Prévention des Handicaps de l'Enfant, 75015 Paris and Pôle de
13 Pédiatrie, Rennes University Hospital, 35203, France

14 4 Department of Paediatric Endocrinology and Diabetology, INSERM CIC-EC5 and Centre de Référence des
15 Maladies Endocriniennes Rares de la Croissance, Robert Debré Hospital and University Paris 7 Denis Diderot,
16 75019, Paris, France

17 4 A complete list of members of the DHCSF study group is given in appendix.
18

19 **Number of words in the abstract: 250**

20 **Number of words in the manuscript: 3398**

21 **Number of figures: 3**

22 **Number of tables: 4 + 1 supplementary table**
23

24 **Corresponding Author:**

25 Prof. Jean-Claude Carel, Paediatric Endocrinology and Diabetology, Hôpital Robert Debré, 48 boulevard

26 Sérurier, 75019 Paris, France. Phone: 33 1 40 03 41 05; Fax: 33 1 40 03 24 29;

27 E-mail: jean-claude.carel@inserm.fr
28
29

30 **Abstract**

31 **Objective.** Neonatal screening for congenital adrenal hyperplasia (CAH) due to 21-
32 hydroxylase deficiency (21OHD) is mainly intended to prevent death due to salt wasting but
33 remains controversial, because of the number of false-positive results and the ease with which
34 most female cases can be identified by virilised genitalia at birth. The aim of this study was
35 to assess the efficiency of the national screening programme for 21OHD.

36 **Design.** Population-based study.

37 **Setting.** National neonatal screening program, paediatric endocrinologists nationwide
38 and reference centre for genotyping.

39 **Participants.** All newborns screened for 21OHD in mainland France between
40 January 1st, 1996 and December 31st 2003.

41 **Outcome Measures.** Screening efficiency indicators, disease severity and
42 contribution of screening to early diagnosis, disease-specific mortality before and during the
43 study period.

44 **Results.** 6,012,798 newborns were screened, 15,407 were considered positive for
45 21OHD and 383 cases were identified, giving a prevalence of 1/15,699 births. The positive
46 predictive value of screening was 2.3% (95% CI, 2.1-2.6), with a sensitivity of 93.5% (90.9-
47 95.9) and a specificity of 99.7%. The false-positive rate was particularly high in preterm
48 infants, for which the positive predictive value was 0.4% (0.2-0.5). Screening allowed
49 clinical diagnosis in 162 of 383 cases (42%), the others being detected clinically or through
50 family history. There was a trend towards declining neonatal mortality due to 21OHD.

51 **Conclusion.** In this large, population-based study, the efficiency of routine 21OHD
52 screening was moderate in neonates born at term and very low in preterm neonates. We
53 recommend the discontinuation of screening, as performed here, in preterm newborns.

54

55 **Introduction**

56 Congenital adrenal hyperplasia (CAH) is a group of autosomal recessive disorders of
57 adrenal steroid biosynthesis ¹⁻³. The commonest form ($\approx 95\%$) is due to 21-hydroxylase
58 deficiency (21OHD). It affects about one child in 15,000 and results in clinical symptoms
59 that vary with the severity of the enzymatic defect. Classical forms include salt-wasting
60 forms (SW), for which there is a high risk of life-threatening adrenal insufficiency during the
61 first month of life, and simple virilising forms (SV). In both cases, female neonates present
62 with markedly virilised external genitalia. Non-classical forms can manifest with
63 hyperandrogenism later in life and do not warrant early recognition through neonatal
64 screening. 3β -hydroxysteroid dehydrogenase deficiency (3β -HSD) is a rare form of CAH that
65 results in the undervirilisation of external genitalia and adrenal insufficiency; it can be
66 detected by 21OHD screening ⁴.

67 21OHD screening is carried out to prevent neonatal death from acute adrenal
68 insufficiency, inaccurate sex assignment in females with complete virilisation and irreversible
69 childhood hyperandrogenism, which may result from incorrect or late diagnosis ^{3,5}. 21OHD
70 fulfils the usual criteria ⁶ for neonatal screening, with its low cost and the availability of a
71 widely applicable test (17-hydroxyprogesterone, 17-OHP, determination) and has been
72 implemented in many Western countries including the USA and some European countries ^{3,7-}
73 ⁹. However, it remains controversial with three main arguments against routine screening: i)
74 the test has a low positive predictive value, with frequent false-positive results in preterm
75 neonates due to cross-reactions with steroids other than 17-OHP ¹⁰, ii) the proportion of cases
76 for which screening really contributes to diagnosis is unclear, as most female cases are easy to
77 detect clinically and salt-wasting is often detected before the screening results are obtained
78 and iii) there is a lack of consensus concerning the 17-OHP threshold to be used, due to
79 changes in 17-OHP distribution with gestational age at birth.

80 In France, 21OHD screening was introduced for all newborns as part of the national
81 screening programme in 1996, after a short pilot feasibility study ¹¹. However, as in many
82 other countries, routine 21OHD screening was never evaluated. The main objective of this
83 study was to evaluate the efficiency of the national French screening program for 21OHD.
84 We retrospectively collected real-life screening data and clinical data for affected neonates, to
85 determine whether screening by the *Association Française pour le Dépistage et la Prévention*
86 *des Handicaps de l'Enfant* (AFDPHE), a national organisation, had facilitated the
87 identification of cases before clinical diagnosis.

88

89 **Methods**

90 **Population studied and data collected**

91 We carried out a retrospective study on all children born in mainland France between
92 January 1st, 1996 and December 31st, 2003. Screening was carried out at 21 regional centres,
93 under the auspices of the *Association Française pour le Dépistage et la Prévention des*
94 *Handicaps de l'Enfant* (AFDPHE)¹². Blood was collected from three-day-old infants on filter
95 paper and 17-OHP concentration was determined by automated time-resolved
96 fluoroimmunoassay (DELFI A®) or RIA. Infants with 17OHP levels above the threshold
97 applied for screening purposes were evaluated further for the diagnosis of 21OHD. We
98 collected data from the regional centres on all newborns for whom 21OHD screening results
99 were considered positive. The data collected included date of birth, gestational age and birth
100 weight, screening and repeat determinations of 17-OHP, assay and threshold used and final
101 conclusions concerning the status of the child: *affected* with CAH (true positive), *unaffected*
102 (false positive) or *deceased*. 17-OHP concentrations are expressed in nmol/l of blood and
103 were converted if necessary (65 pg/spot = 80 nmol/l of blood). The threshold applied was
104 that recommended nationally by the AFDPHE, but was modified slightly at different times
105 and in different regions, based on the local distribution of 17-OHP levels. We collected
106 additional data from the medical records of affected children, concerning sex, date, weight
107 and plasma sodium concentration at diagnosis, genital abnormalities classified as described by
108 Prader¹ and *CYP21A2* genotyping results, classified as classical salt wasting (SW), classical
109 simple virilising (SV) or non classical forms^{13, 14}. If genotyping results were not available or
110 not informative (n = 2) due to the detection of mutations with unknown functional
111 repercussions, patients were classified as a function of the clinical data, leaving only one
112 unclassified patient, who was then arbitrarily classified as affected with SW CAH. Weight at
113 diagnosis was expressed as a percentage of expected weight at a given age, based on birth

114 weight and the expected 1% gain in weight per day after day 8¹⁵. The distribution of the
115 gestational ages of true negatives was derived from reference values published annually in
116 France (DRESS 2001).

117 We searched for false negatives (FN) detected before March 2010, which is at least six
118 years after the birth of the last child studied, using five data sources: 1) regional screening
119 centres notified of FN cases by physicians; 2) mail and e-mail surveys of all paediatric
120 endocrinologists registered with the national society or treating children with CAH; 3) the
121 French reference centre for CAH genotyping in Lyon and another molecular biology
122 laboratory performing CAH genotyping and 4) the Centre for Epidemiology Medical Causes
123 of Death database (CépiDc, INSERM), in which we looked for children dying from causes
124 corresponding to International Classification of Diseases (ICD) 9 and 10 codes 255.2, 255.4,
125 E25, E27.4 (adrenogenital disorders, other and unspecified adrenocortical insufficiency).

126 **Statistical analysis**

127 We calculated the sensitivity, specificity and predictive values of the screening test,
128 with 95% confidence intervals (CI) for preterm neonates born before 37 weeks of gestational
129 age, for term newborns, and for both considered together. We classified the contribution of
130 CAH screening to the diagnosis of true positives as follows: screening was considered *useful*
131 if it led to the diagnosis of classical 21OHD (SW or SV forms) or 3 β -HSD deficiency that was
132 not suspected clinically because there were no symptoms or because the symptoms and signs
133 (genital abnormalities, dehydration) had not been recognised; screening was considered *not*
134 *useful* if CAH was diagnosed before the results of screening became available (on the basis of
135 family history, prenatal diagnosis or neonatal systematic examination). Screening was also
136 considered *not useful* for false negative cases of classical CAH and for children with positive
137 screening results diagnosed with non-classical forms of CAH.

138 The relationship between gestational age at birth and 17-OHP concentration was
139 studied by linear regression analysis in a sample of 10,523 preterm neonates born before 37
140 weeks of amenorrhoea selected from the infants testing positive. The values for 17-OHP
141 concentration were not normally distributed and a natural logarithm transformation was
142 therefore applied. Goodness of fit (R^2) was calculated for various linear regression models, to
143 identify the factor best predicting 17-OHP concentration: gestational age or birth weight.
144 Linear regression models were constructed for the imputation of missing data for term or birth
145 weight.

146 Mortality rates for children under the age of one year were calculated between 1979
147 and 2007, from CepiDc data. Changes in mortality rate over time were assessed by Poisson
148 regression analysis. We looked for a possible change in slope after 1996 (the year in which
149 the screening programme was generalised), by looking for an interaction between “year”,
150 considered as a continuous variable, and “before/after screening introduction” considered as a
151 dichotomous variable.

152 All analyses were performed with SAS 9.2 software (SAS Institute, Cary, NC, USA).
153 The study was approved by the CCTIRS and CNIL and was conducted in accordance with
154 French legislation.

155

156 **Results**

157 During the eight-year study period, 6,012,798 screening tests for 21OHD were
158 performed on children born in mainland France (Figure 1, Table 1 and supplementary Table
159 1). The laboratory methods for 17-OHP determinations and their thresholds are shown in
160 supplementary Table 1. Neonatal screening tests were positive for 15,407 newborns, with
161 370 considered *affected*, 11,324 considered *unaffected* and no conclusion reached by the
162 screening centres for 3,132. For 1,814 infants, the conclusion was discordant with the last
163 recorded 17-OHP concentration (n = 338 considered *unaffected* with a last 17-OHP
164 determination considered positive and n = 1476 considered *unaffected* without the recorded
165 monitoring of 17-OHP concentrations). 581 children were identified as *deceased* : most of
166 these children were preterm and, in all cases, the death of the child was considered by clinical
167 centres to be unrelated to 21OHD (Figure 1). Most of the newborns with positive results for
168 21OHD screening were born before term (91% of those for whom data were available). The
169 median day for filter paper sampling was day 4, although the screening protocol called for
170 sampling on day 3. Of the 370 newborns considered to be affected, 358 had a classical form
171 of 21OHD (n = 354) or 3 β HSD (n = 4) deficiency and 12 had a non-classical form of CAH.

172 The median age at diagnosis of CAH was seven days (Table 2). Weight loss was
173 severe (>10% of expected body weight) in 19% of those for whom data were available, and
174 plasma sodium concentration was below 130 mmol/l in 18% of the infants. Screening was
175 useful for diagnosis in 162 of the 358 children with classical CAH and positive screening
176 results, mostly males with the SW form (n = 106/162). Screening results were positive but
177 not useful for diagnosis in 74 children with a family history of 21OHD and in 96 girls with
178 genital abnormalities detected during neonatal examination. In addition, screening results
179 were positive in 13 boys with classical 21OHD who were diagnosed clinically before the
180 screening results became available. Of interest, among the 38 premature babies with positive

181 screens, screening was useful to the diagnosis in only 13 among whom only 6 had a SW form.
182 We identified 25 children as false negative for 21OHD screening: 23 were reported by
183 genotyping laboratories, and 20 of these cases were also reported by the screening centres,
184 with two reported by the CépiDC. Most of the false negatives (16/25) had SV forms (Figure
185 1).

186 Altogether, the incidence of classical 21OHD (SW and SV forms) and 3 β -HSD
187 deficiency in France between 0 and 1 year was 0.78/10⁵ births/year, with a 95% CI of [0.70 –
188 0.86] and the prevalence was 1/15699 births (95%CI, 1/17445 - 1/14269) (including false
189 negative subjects in their birth cohort). The sensitivity of screening was 93.5% with an
190 overall positive predictive value of 2.3% (Table 3). Sensitivity was higher for SW 21OHD
191 (96.9%, 95%CI, 94.8-98.9) than for SV 21OHD (82.8%, 95%CI, 75.1-90.5). Most false-
192 positive screening test results were obtained for preterm newborns, for which the positive
193 predictive value of screening was only 0.36%, whereas that for term newborns was 30.4%.
194 We investigated whether adjustment of the 17-OHP threshold would have improved screening
195 efficiency in preterm newborns, by calculating linear regression models of (positive) 17-OHP
196 levels on filter paper. Gestational age accounted for 9.5% and weight accounted for 7.4% of
197 the variance (R²) of 17-OHP concentration. Adding polynomials and assay techniques
198 increased the R² to 10%. Figure 2 illustrates the difficulty of establishing threshold values
199 based on gestational age.

200 As the primary objective of 21OHD screening is to prevent the death of newborns, we
201 analysed 21OHD-related mortality in France from 1979 to 2007, a 29-year period including
202 the year in which 21OHD screening was introduced. Twenty-one children under the age of
203 one year were classified with an underlying cause of death due to adrenogenital disorders,
204 other and unspecified adrenocortical insufficiency (Figure 3). There was a significant
205 (p=0.002) trend towards a decrease in specific mortality rate during this period, with most of

206 this decrease occurring in 1991 to 1995, before the generalisation of screening. Thus, neither
207 screening itself (yes/no) nor the interaction of screening and time was associated with specific
208 mortality rate ($p=0.31$ and 0.31 , respectively).

209

210 **Discussion**

211 With the inclusion of 6,012,798 newborns screened in mainland France between 1996
212 and 2003, this study is by far the largest to date to assess neonatal screening for 21OHD with
213 particular emphasis on its contribution to early diagnosis. We found that sensitivity was good
214 (93%), but that the positive predictive value of screening was low (2.3%), although it
215 improved markedly if we considered only term newborns (30.4%). Screening results
216 contributed to diagnosis in 42% of the cases. Moreover, the large number of infants for which
217 no conclusion was drawn raises questions about the practical organisation of 21OHD
218 screening, due to the large number of false positives.

219 Table 4 summarises published data from previous population studies, making
220 comparisons with our results possible. The positive predictive values reported in these studies
221 were similar in most cases, with the exception of the Swiss study (positive predictive value of
222 50%), which presented results for a second determination of 17-OHP on filter paper, rather
223 than those for the primary screening, as in most studies. Unlike previous studies, we took
224 gestational age into account, and we found that screening efficiency differed considerably
225 between term and preterm newborns. Among preterm newborns, there were almost 277 false
226 positives for each case of 21OHD discovered, whereas there were only two to three false
227 positives for each case for term newborns. These difficulties arise from the low specificity of
228 immunological assay techniques for determining levels of 17-OHP in preterm newborns due
229 to high plasma concentrations of steroids other than 17-OHP that cross-react in the assays
230 (with sulphated metabolites), generating false-positive results ¹⁶⁻¹⁸. Some countries have
231 adopted variable threshold values based on gestational age (Table 4), but our study shows that
232 there is a large overlap of 17-OHP levels between affected and unaffected preterm newborns
233 and that increasing the threshold level in this particular population would result in a loss of
234 sensitivity. One possible alternative is the use of tandem mass spectrometry as a second line

235 test to improve the positive predictive value of screening ^{10,19}. These techniques were recently
236 recommended in the Endocrine Society guidelines ³, but they are costly, not widely available
237 for population screening and require thorough evaluation, including cost-benefit analyses.

238 Although 21OHD screening correctly identified 93% of cases, its impact on diagnosis
239 was much smaller, as it contributed to early diagnosis in 45 to 50% of the children identified,
240 corresponding to about 20 children per year in France or an incidence of 2.66/10⁵/year. The
241 main reasons for this minor contribution are that girls with classical 21OHD are readily
242 identified during neonatal paediatric examination and CAH is an autosomal recessive
243 disorder, making prenatal or neonatal diagnosis more likely in families with an index case. In
244 addition, in a small proportion of boys (9/153) with SW forms, adrenal crisis occurred before
245 screening results became available and the children were correctly managed based on their
246 clinical presentation.

247 Screening for 21OHD is designed principally to decrease neonatal disease-specific
248 mortality. A decrease in specific mortality has been observed over the last three decades, but
249 the timing of this decrease suggests that it was due to improvements in paediatric care rather
250 than to the introduction of screening. The probability of death due to neonatal adrenal crisis
251 in the absence of screening is widely debated and has been reported to vary from 0 to 4% of
252 patients with SW 21OHD in populations with high standards of clinical awareness and care
253 for 21OHD ⁷. In our study population of 285 children with salt-wasting 21OHD born
254 between 1996 and 2003 (276 true positives and 9 false negatives), using 4% as an estimate
255 suggests that 11.5 neonatal deaths would have been expected in the absence of screening, a
256 figure to compare to 3 deaths observed during the first year of life. In addition to preventing
257 mortality, screening for 21OHD can prevent inaccurate gender assignment and irreversible
258 childhood hyperandrogenism. In our study, inaccurate gender assignment was not made in
259 the 5 fully virilised females (Prader stage V) but screening allowed the identification of 47/77

260 patients with a SV form confirming the value of screening to detect the 21OHD before the
261 appearance of severe hyperandrogenism.

262 Our findings also show that the organisation of screening, as currently conducted in
263 France (and possibly elsewhere), was not satisfactory. Screening centres encountered major
264 operational difficulties with follow-up of the large number of positive tests and no conclusion
265 about status was reached in many cases, raising medical, ethical and responsibility issues.
266 Questions remain concerning the fate of 307 children for whom successive assays remained
267 above the threshold value but for whom no further follow-up data were obtained. Data for
268 weight and gestational age were also frequently missing, although this information makes it
269 easier to interpret the assay results and should therefore be collected.

270 Our study was subject to several limitations. The apparent lack of impact of screening
271 on mortality is likely due to insufficient statistical power, given the limited numbers in the
272 mortality analysis (21 deaths in 29 years), the unmeasured effect of pilot 21OHD screening
273 programs during the 1990-1995 period and more importantly undiagnosed 21OHD-related
274 deaths in the period before screening as documented in central Europe ²⁰. In addition,
275 inclusion in our analysis of misclassified adrenal diseases other than CAH might have
276 obscured the effect of screening on mortality. However, the study had sufficient power (80%)
277 to detect a decrease in specific mortality of 65% or more after screening implementation. We
278 could not trace with precision the reasons for which false-negative cases were missed (17-
279 OHP concentration below the threshold, positive test not taken into account correctly, other
280 reason). False-negative cases may have been underestimated for several reasons, including
281 lack of ascertainment if cases were detected later in life and not followed by paediatric
282 endocrinologists or subjected to molecular analyses. We were also unable to determine the
283 number of unnecessary visits and laboratory investigations for children eventually found not
284 to have 21OHD, or their impact on the anxiety of the parents. Missing data on gestational age

285 may have resulted in an overestimation of the sensitivity of the screening, particularly in
286 subgroup analysis (preterm and term newborns). One further limitation is that data on
287 neonatal screening from 1996-2003 are only presented in 2011, at a time when they might be
288 considered less timely. This apparent delay results from long and tedious data collection and
289 monitoring and from the need to wait several years in order to be able to identify FN SV
290 forms since the diagnosis is made as late as 5 or 6 years in some cases. Indeed, we searched
291 for FN cases in 2010, at a time when the youngest child in the cohort was older than 6 years,
292 which is an additional strength of our study.

293 In France, the budget per screened infant (for 21-OHD only) is €1.23 in the absence of
294 follow-up (US\$ 1.7), corresponding to approximately €924,500/year. As ≈ 20 cases of
295 21OHD are identified by screening each year, this budget corresponds to approximately
296 €50,000 (US\$ 70,000) per case. In the US, the cost per screened infant without follow-up has
297 been estimated at US\$ 2.3 to 6.0 and applying these figures to our data would result in
298 estimates of US\$ 95,000 to 245,000 per case²¹. Assumptions concerning the number of
299 potential deaths among these cases and a complete medical economic analysis of indirect
300 costs might allow a full evaluation of the cost per life-year saved²¹.

301 Overall, we found that neonatal 21OHD screening was efficient in term newborns,
302 with a variable impact on clinical management given that most affected female newborns are
303 easy to identify without screening. By contrast, the efficiency of screening was very low in
304 preterm newborns, resulting in large numbers of false positives, flooding the system and
305 leading to its dysfunction and leading to the identification of 6 cases of potentially lethal salt-
306 wasting 21OHD among more than 10,000 positives in 8 years. Improved organisation might
307 have prevented this dysfunction and allowed a comprehensive follow-up of all positive cases.
308 However, a decrease of the false positive cases is necessary to improve efficiency.

309 So, what recommendations should be made based on our results? We recommend that
310 newborn screening for 21OHD is continued for term newborns in areas in which it is already
311 performed and that careful consideration is given to its implementation in areas in which this
312 is not the case. By contrast, we recommend that 21OHD screening, as performed in this
313 study, should not be carried out for preterm newborns since the positive predictive value of
314 the test is very low and most preterm newborns are subject to careful paediatric care that
315 should ensure that incipient salt wasting adrenal crises are readily recognised. In France, the
316 national neonatal screening organisation and representatives of professional organisations in
317 neonatology, paediatric endocrinology and rare endocrine diseases are currently discussing
318 how to improve the national screening program. Our improved program, as well as others
319 around the world will have to be carefully evaluated.

320

321 **Acknowledgment**

322 The authors wish to acknowledge the role of Unité de Recherche Clinique Paris-
323 Cochin and in particular Christelle Auger and Sandra Colas. Cécile Klochendler, Sandrine
324 Borderie and Pr Jean-Marc Treluyer from Unité de Recherche Clinique Paris-Cochin are also
325 acknowledged. Louise Montagne, Claire Jeandel, Kasia Marchal, Candace Bensignor, Chaza
326 Fallouh and the Centres of Clinical Investigation of Lyon and Bordeaux also participated in
327 data collection.

328

329 **Authors' contributions statement**

330 Jean-Claude Carel, Joel Coste and Yves Morel conceived and conducted the study,
331 and analysed the data. Emmanuel Ecosse established the database and some of the statistical
332 analyses. Bénédicte Coulm participated in the conduction of the study analyzed the data.
333 Bénédicte Coulm, Jean-Claude Carel and Joel Coste had full access to all the data in the study
334 and take responsibility for the integrity of the data and the accuracy of the data analysis and
335 wrote the paper. Yves Morel and Véronique Tardy performed the molecular analysis and the
336 classification of patients. Michel Roussey participated in the conduction of the study and
337 contributed to data analysis. All authors participated in the elaboration of the manuscript and
338 commented on it. The members of DHCSF study group are to be considered as co-authors of
339 the manuscript given their involvement in the elaboration of the protocol, data collection or
340 follow-up of patients. The final version of the English text was edited by Julie Sappa (Alex
341 Edelman & Associates, Scientific Editors).

342

343 **Source of funding statement**

344 The study was funded by the French Ministry of Health "PHRC, Programme
345 Hospitalier de Recherche Clinique" grant number AOM03048, by the French National Health
346 Insurance "CNAM, Caisse Nationale d'Assurance Maladie" and by the French National
347 Neonatal Screening Programme Organisation (AFDPHE, Association Française pour le
348 Dépistage et la Prévention des Handicaps de l'Enfant). The PHRC and CNAM did not
349 participate in the study or its interpretation. AFDPHE was involved in the conduction of the
350 study and its president (Michel Roussey) is co-author of the paper.

351

352 **Conflict of interest statement**

353 No conflict of interest to disclose.

354 **DHCSF study group.**

355 *Investigators who participated in the elaboration and conduction of the study:*

356 Dr Claire Bouvattier, Kremlin-Bicêtre; Prof. Régis Coutant, Anger; Dr Luc Desfrère, Colombes; Prof. Jean-
357 Pierre Farriaux, Lille; Prof. Juliane Léger, Paris; Prof. Marc Nicolino, Lyon; Prof. Michel Polak, Paris; Dr
358 Marie-Charles Raux-Demay, Paris; Prof. Maité Tauber, Toulouse

359 *Members of regional screening associations who organized and conducted the study:*

360 Alsace, Ms Lehr, Prof. Juif; Angers, Dr Berthelot, Prof. Ginies; Aquitaine, Ms Dubeau, Prof. Lacombe;
361 Auvergne, Ms Roche, Prof. Demeocq.; Bourgogne, Ms Jorrot Ms Frigère, Prof. Huet; Bretagne, Ms Rosselin,
362 Prof. Roussey; Centre, Ms Vicente, Prof. Besnard; Champagne-Ardenne, Ms Brissiaud, Dr Garnotel; Franche-
363 Comte, Ms Gaulard, Dr Schirrer; Ile de France, Ms Gauthereau, Prof. Czernichow; Languedoc-Roussillon, Ms
364 Saint-Dizier, Prof. Sarda; Limousin, Ms Chebroux, Dr Constanty; Lorraine, , Prof. Leheup; Midi-Pyrénées, Ms
365 Rue, Ms Fauvel, Prof. Olives; Nantes, Ms Mirallié, Prof. Barriere; Nord-Pas-de-Calais, Ms Paux, Prof. Farriaux;
366 Normandie, , Dr Travert; Paca, Dr Simonin, Prof. Sarles; Picardie, Ms Ducrout, Prof. Boudailliez; Rhone-Alpes
367 Grenoble, Ms Gachet, Ms Picard, Prof. Debillon; Rhone-Alpes Lyon, Ms Debiase, Ms Dorche, Prof.
368 Gaucherand, Prof. Debillon; CH de Chambéry, Dr Vercherat.

369 *Laboratories who conducted the genotyping*

370 Prof. Morel, Dr Tardy, Lyon; Dr Bellane-Chantelot, Paris-Pitié Salpêtrière; Prof. Clauser, Paris-Cochin.

371 *Physicians who participated in the evaluation and management of screened newborns*

372 Dr Adiceam, Aix-en-Provence; Dr Boudailliez, Amiens; Prof. Coutant, Angers; Dr Duquesne, Auch; Dr
373 Vedrenne, Avon; Dr Pincemaille, Bastia; Dr Carre, Bayonne; Dr Bertrand, Besançon; Dr Dieckmann, Blois; Dr
374 Sauvion, Bondy; Dr Barat, Bordeaux; Dr Llanas, Bordeaux; Dr Sarlangue, Bordeaux; Dr Riviere, Bourges; Dr
375 Giroux, Brest; Dr Metz, Brest; Dr Cuvelier, Calais; Dr Elchardus, Charleville-Mézières; Dr Ezzeddine,
376 Charleville-Mézières; Dr Phan, Chartres; Dr Benchekroun, Châteauroux; Dr Carla, Clermont-Ferrand; Dr
377 Pennerath, Colmar; Dr Vervel, Compiègne; Dr Chevrel, Dax; Dr Jeannot, Dieppe; Dr de Monleon, Dijon; Dr
378 Gounot, Dijon; Prof. Huet, Dijon; Dr Loeuille, Dunkerque; Dr Dulucq, Epinal; Dr Kozisek, Flers; Dr Al-Issa,
379 Fontainebleau; Dr Tommasi, Grasse; Dr Dupuis, Grenoble; Dr Pigeon, Hazebrouck; Dr Brossier, La-Roche-sur-
380 Yon; Prof. Brauner, Le-Kremlin-Bicêtre; Dr Bonardi, Le-Mans; Dr Maxaud, Le-Mans; Dr Souto, Le-Mans; Dr
381 Chauvet, Lens; Dr Guemas, Lens; Dr Boulard, Libourne; Dr Cartigny-Maciejewski, Lille; Prof. Lienhardt,
382 Limoges; Dr Ribault, Lisieux; Dr Naud-Saudreau, Lorient; Dr Berlier, Lyon; Prof. Chatelain, Lyon; Dr Nicolino,
383 Lyon; Dr Simonin, Marseille; Dr Mathivon, Meaux; Dr Pignol, Mont-De-Marsan; Dr Arzim, Montélimar; Dr
384 Jeandel, Montpellier; Prof. Sultan, Montpellier; Dr Benoit, Mulhouse; Dr Baron, Nantes; Dr Ramos, Nantes; Dr
385 Baechler-Sadoul, Nice; Dr Wagner, Nice; Dr Zelinsky, Niort; Dr Monceaux, Orléans; Dr Moretti, Orsay; Dr
386 Beaussac, Paimpol; Dr Cabrol, Paris; Prof. Leger, Paris; Dr Morel-Bouvattier, Paris; Prof. Polak, Paris; Dr
387 Pradeaux, Perigueux; Dr Barba, Pessac; Dr Puel, Pessac; Dr Crosnier, Poissy; Dr Sarda, Pontoise; Dr Queinnec,
388 Quimper; Dr Sulmont, Reims; Dr de Kerdanet, Rennes; Dr Jeannoel, Roanne; Dr Ythier, Roubaix; Prof. Mallet,
389 Rouen; Dr Idres, Saint-Brieux; Dr Raynaud-Ravni, Saint-Etienne; Dr Chouraki, Saint-Quentin; Dr Garandeau,
390 StDenis-LaRéunion; Dr Soskin, Strasbourg; Dr Petrus, Tarbes; Dr Feldmann, Thionville; Dr Lambert-Leonardi,
391 Thionville; Dr Jesuran-Perelroizen, Toulouse; Prof. Tauber, Toulouse; Dr Despert, Tours; Dr Ceccato, Tresses;
392 Dr Jullien, Troyes; Dr Ninot, Troyes; Dr Soulier, Tulle; Dr Leheup, Vandoeuvres LesNancy; Dr Goldfarb,
393 Vannes; Dr Goumy, Vichy; Dr Rebaud, Villefranche-Sur-Saone; Dr Roubin, Villeneuve-Sur-Lot.

394

395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453

References

1. White PC, Speiser PW. Congenital adrenal hyperplasia due to 21-hydroxylase deficiency. *Endocr Rev.* 2000 Jun 2000;21(3):245-291.
2. Merke DP, Bornstein SR. Congenital adrenal hyperplasia. *Lancet.* Jun 18-24 2005;365(9477):2125-2136.
3. Speiser PW, Azziz R, Baskin LS, et al. Congenital adrenal hyperplasia due to steroid 21-hydroxylase deficiency: an Endocrine Society clinical practice guideline. *J Clin Endocrinol Metab.* Sep 2010;95(9):4133-4160.
4. Johannsen TH, Mallet D, Dige-Petersen H, et al. Delayed diagnosis of congenital adrenal hyperplasia with salt wasting due to type II 3beta-hydroxysteroid dehydrogenase deficiency. *J Clin Endocrinol Metab.* Apr 2005;90(4):2076-2080.
5. Pang SY, Wallace MA, Hofman L, et al. Worldwide experience in newborn screening for classical congenital adrenal hyperplasia due to 21-hydroxylase deficiency. *Pediatrics.* Jun 1988;81(6):866-874.
6. Wilson JM. The evaluation of the worth of early disease detection. *J R Coll Gen Pract.* Nov 1968;16 Suppl 2:48-57.
7. Grosse SD, Van Vliet G. How many deaths can be prevented by newborn screening for congenital adrenal hyperplasia? *Horm Res.* 2007;67(6):284-291.
8. National newborn screening and genetics resource center. http://genes-r-us.uthscsa.edu/resources/newborn/newborn_menu.htm.
9. Loeber JG. Neonatal screening in Europe; the situation in 2004. *J Inherit Metab Dis.* Aug 2007;30(4):430-438.
10. Janzen N, Peter M, Sander S, et al. Newborn screening for congenital adrenal hyperplasia: additional steroid profile using liquid chromatography-tandem mass spectrometry. *J Clin Endocrinol Metab.* Jul 2007;92(7):2581-2589.
11. Cartigny-Maciejewski M, Guilley N, Vanderbecken S, et al. Neonatal screening of congenital adrenal hyperplasia due to 21-hydroxylase deficiency: Lille experience 1980-1996. *Arch Pediatr.* 1999 Feb 1999;6(2):151-158.
12. Dhondt JL, Farriaux JP. The wonderful history of neonatal screening. *Ann Biol Clin (Paris).* May-Jun 2000;58(3):267-276.
13. Morel Y, Miller WL. Clinical and molecular genetics of congenital adrenal hyperplasia due to 21-hydroxylase deficiency. In: Harris H, Hirschhorn K, eds. *Advances in human genetics.* 20 ed. New York: Plenum Press; 1991:1-68.
14. Tardy V, Menassa R, Sulmont V, et al. Phenotype-genotype correlations of 13 rare CYP21A2 mutations detected in 46 patients affected with 21-hydroxylase deficiency and in one carrier. *J Clin Endocrinol Metab.* Mar 2010;95(3):1288-1300.
15. van Dommelen P, van Wouwe JP, Breuning-Boers JM, van Buuren S, Verkerk PH. Reference chart for relative weight change to detect hypernatraemic dehydration. *Arch Dis Child.* Jun 2007;92(6):490-494.
16. Nordenstrom A, Wedell A, Hagenfeldt L, Marcus C, Larsson A. Neonatal screening for congenital adrenal hyperplasia: 17-hydroxyprogesterone levels and CYP21 genotypes in preterm infants. *Pediatrics.* 2001 Oct 2001;108(4):E68.
17. Wong T, Shackleton CH, Covey TR, Ellis G. Identification of the steroids in neonatal plasma that interfere with 17 alpha-hydroxyprogesterone radioimmunoassays. *Clin Chem.* Sep 1992;38(9):1830-1837.
18. al Saedi S, Dean H, Dent W, Cronin C. Reference ranges for serum cortisol and 17-hydroxyprogesterone levels in preterm infants. *J Pediatr.* Jun 1995;126(6):985-987.
19. Minutti CZ, Lacey JM, Magera MJ, et al. Steroid profiling by tandem mass spectrometry improves the positive predictive value of newborn screening for congenital adrenal hyperplasia. *J Clin Endocrinol Metab.* Aug 2004;89(8):3687-3693.
20. Kovacs J, Votava F, Heinze G, et al. Lessons from 30 years of clinical diagnosis and treatment of congenital adrenal hyperplasia in five middle European countries. *J Clin Endocrinol Metab.* Jul 2001;86(7):2958-2964.
21. Yoo BK, Grosse SD. The cost effectiveness of screening newborns for congenital adrenal hyperplasia. *Public Health Genomics.* 2009;12(2):67-72.
22. Therrell BLJ, Berenbaum SA, Manter-Kapanke V, et al. Results of screening 1.9 million Texas newborns for 21-hydroxylase-deficient congenital adrenal hyperplasia. *Pediatrics.* 1998 Apr 1998;101(4 Pt 1):583-590.
23. Steigert M, Schoenle EJ, Biason-Lauber A, Torresani T. High reliability of neonatal screening for congenital adrenal hyperplasia in Switzerland. *J Clin Endocrinol Metab.* Sep 2002;87(9):4106-4110.
24. Van der Kamp HJ, Noordam K, Elvers B, Van Baarle M, Otten BJ, Verkerk PH. Newborn screening for congenital adrenal hyperplasia in the Netherlands. *Pediatrics.* 2001 Dec 2001;108(6):1320-1324.

454
455
456
457
458
459

25. Balsamo A, Cacciari E, Piazzì S, et al. Congenital adrenal hyperplasia: neonatal mass screening compared with clinical diagnosis only in the Emilia-Romagna region of Italy, 1980- 1995. *Pediatrics*. 1996 Sep 1996;98(3 Pt 1):362-367.

460 **Legends for figures**

461

462 **Figure 1. Results of neonatal screening for 21-hydroxylase deficiency in France 1996-2003**

463

464

465 **Figure 2. 17-OHP concentration (nmol/l) at neonatal screening in affected and unaffected (false-**
466 **positive) children as a function of gestational age**

467 Individual values for affected children are shown as red dots ♦ and those for false positives are

468 shown as green dots * .

469

470

471 **Figure 3. Specific mortality rates due to adrenogenital disorders, other and unspecified**

472 **adrenocortical insufficiency during the first year of life in France, 1979-2007.**

473

474

475

476 **Table 1 – Principal characteristics of newborns with positive screening results for CAH**
 477 **in mainland France between 1996 and 2003**
 478

N = 15,407	
Sex	
Male	9,031 (59)
Female	6,218 (41)
Gestational age (WA) ^{†*}	
Preterm (<37 WA)	10,563 (68)
Term (≥37 WA)	1,058 (7)
Missing data	3,786 (25)
Birth weight (g) ^{†*}	
	1,490 (1,005 – 2,090)
Age at screening (days) [†]	
	4 (3-5)
17-OHP screening result (nmol/l) [†]	
	80 (65-109)
17-OHP measurement method used	
Radioimmunoassay	8,457 (55)
DELFI [®]	6,945 (45)

479
 480 Data are presented as numbers (percentages) unless otherwise stated; †, median, interquartile range; WA: weeks
 481 of amenorrhoea; * imputations of gestational age from sex and birth weight: n = 998, imputations of birth weight
 482 from sex and gestational age: n = 2653; † missing data: sex n = 158; age at screening n = 1117; 17-OHP
 483 screening result n = 72; 17-OHP measurement method n = 5.

484

485

486 **Table 2 – Characteristics of affected newborns with CAH due to classical 21OHD or 3β-**
 487 **HSD detected by screening in mainland France between 1996 and 2003**

	True positives (n = 358)
Sex	
Boys	205 (57)
Girls	153 (43)
Gestational age (WA) [†]	39 (38-40)
Preterm (<37 WA)	38 (10.7)
Term (≥37 WA)	318 (89.3)
Birth weight (g) [†]	3,370 (2,980 – 3,680)
Age at screening (days) [†]	3 (3-4)
Age at diagnosis (days) [†]	7 (1-10)
Contribution of screening to the diagnosis of CAH, n (M/F)	
Useful	162 (137/25)
Salt-wasting 21OHD	114 (106/8)
Simple virilising 21OHD	47 (30/17)
3β-HSD	1 (1/0)
Not useful*	196 (68/128)
Clinical diagnosis before screening results	109 (13/96)
Salt-wasting 21OHD	99 (9/90)
Simple virilising 21OHD	8 (2/6)
3β-HSD	2 (2/0)
Prenatal diagnosis or family history	74 (46/28)
Salt-wasting 21OHD	54 (33/21)
Simple virilising 21OHD	20 (13/7)
3β-HSD	0 (0/0)
Information on usefulness unavailable	13 (9/4)
Salt-wasting 21OHD	10 (6/4)
Simple virilising 21OHD	2 (2/0)
3β-HSD	1(1/0)
Plasma sodium concentration at diagnosis (nmol/l) [†]	
≥135	177 (78M/99F)
130-135	80 (61M/19F)
<130	56 (48M/8F)
Relative weight change at diagnosis (% of expected) [†]	
≥0%	42 (21/21)
[0; -5%]	62 (34/28)
[-5; -10%]	76 (60/16)
<-10%	42 (39/3)

488

489 Data are presented as numbers (percentages) unless otherwise stated; †, median, interquartile range; WA: weeks
 490 of amenorrhoea; SD: Standard deviation; missing data: gestational age, n=2; age at screening, n = 28; age at
 491 diagnosis, n = 38; plasma sodium concentration, n = 47; relative weight change at diagnosis, n = 138. The 25
 492 false-negative cases (11 boys and 14 girls), for whom screening was not useful, are not included in this table.
 493

494 **Table 3. Efficiency of 21OHD screening as a function of gestational age at birth**

495 *Table 3a: Raw data*

	Screening	Affected	Unaffected	Total
All newborns	Positive	358	15,049	15,407
	Negative	25	5,997,366	5,997,391
	Total	383	6,012,415	6,012,798
Term newborns (≥37 WA)	Positive	318	740	1058
	Negative	21	5,578,196	5,578,217
	Total	339	5,578,936	5,579,275
Preterm newborns (<37 WA)	Positive	38	10,524	10,562
	Negative	2	422,959	422,961
	Total	40	433,383	433,523

496

497 *Table 3b: Efficiency calculations*

All newborns	Positive predictive value	2.3 [2.1-2.6]	Specificity	99.7 [99.7 - 99.7]
	Negative predictive value	99.9 [99.9 - 99.9]	Sensitivity	93.5 [90.9 - 95.9]
Term newborns (≥37 WA)	Positive predictive value	30.1 [27.3-32.8]	Specificity	99.9 [99.9-99.9]
	Negative predictive value	99.9 [99.9 ; 99.9]	Sensitivity	93.8 [91.2-96.4]
Preterm newborns (<37 WA)	Positive predictive value	0.4 [0.2-0.5]	Specificity	97.6 [99.5-97.6]
	Negative predictive value	99.9 [99.9 ; 99.9]	Sensitivity	95 [83.1 -99.4]

498

499 Data are expressed in % [95% confidence interval]; efficiency in term and preterm newborns was calculated for
 500 those without missing data for gestational age at birth (11,620/15,407).

501

502 **Table 4. Efficiency of 21OHD screening in published studies**

Reference	Country	Number of newborns	17-OHP threshold (nmol/l)	Variable 17-OHP threshold with term	Sensitivity (%)	Positive predictive value (%)
²²	USA (Texas)	1,936,998	123	yes	86	NA
¹¹	France	408,138	36 to 60	no	89	2.1
²³	Switzerland	333,221	30 to 90	yes	97	50.0
²⁴	Netherlands	176,684	60	yes	100	5.9
²⁵	Italy	128,330	36	yes	NA	1.9
This study	France	6,012,798	40 to 100	no	93	2.3

503

504 Population-based studies with a sample size >100 000 were selected

505

506 **Supplementary Table 1: Thresholds and laboratory methods used for 17-OHP**
 507 **determination**
 508

Methods and thresholds	Number of newborns who screened positive , n (%)
Delfia®*	6,945 (45.1)
40 nmol/l	595
50 nmol/l	3,209
60 nmol/l	1,416
70 nmol/l	1,725
RIA*	8,457 (54.9)
50 nmol/l	1,572
60 nmol/l	4,050
70 nmol/l	1,209
80 nmol/l	87
100 nmol/l	78
50 pg/spot	1,124
60 pg/spot	34
70 pg/spot	303

509
 510 *Delfia®, dissociation-enhanced lanthanide fluorescence immunoassay; RIA, radioimmunoassay; missing data:
 511 Laboratory methods, n = 5, thresholds n = 5.
 512

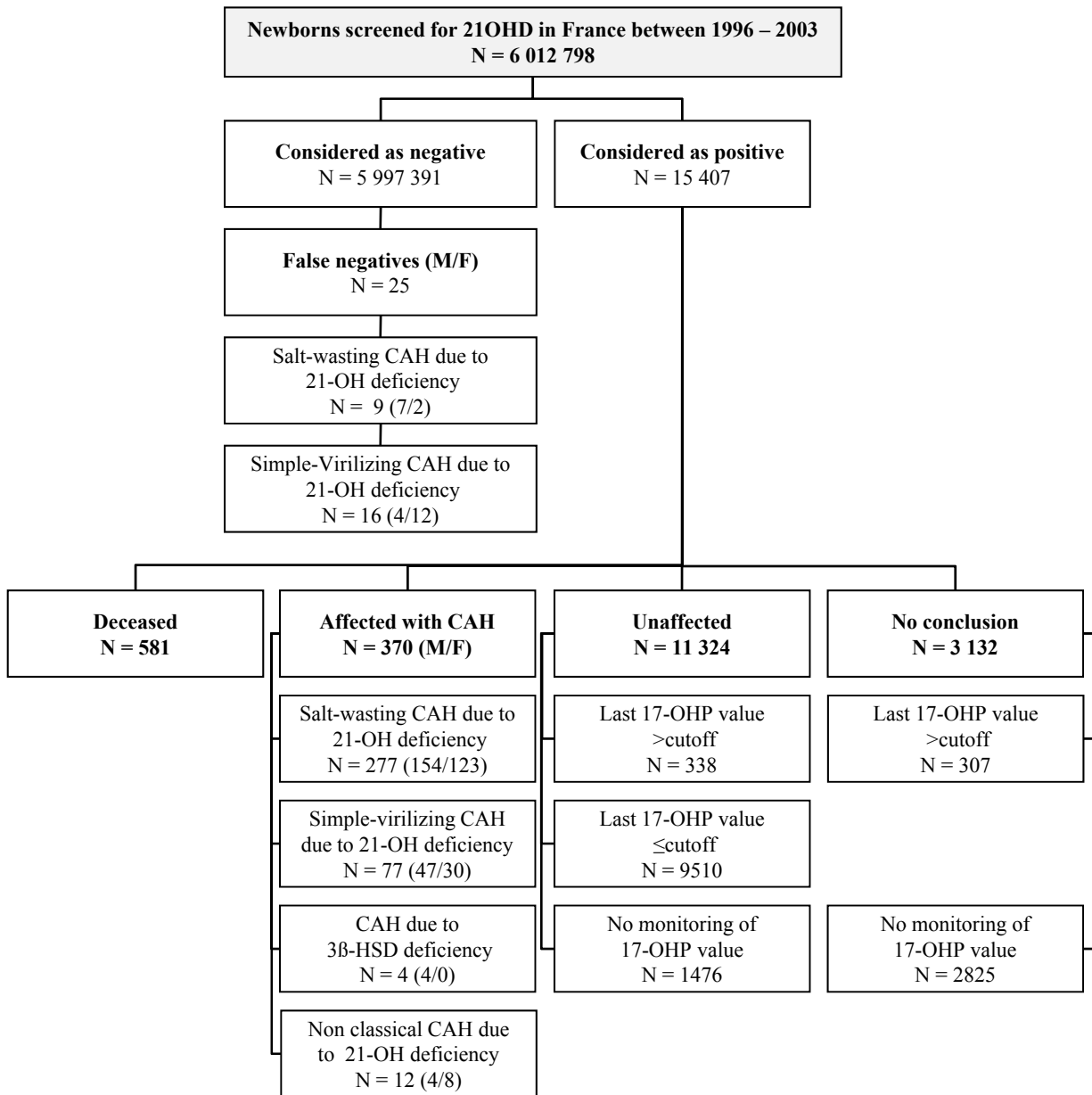


Figure 1

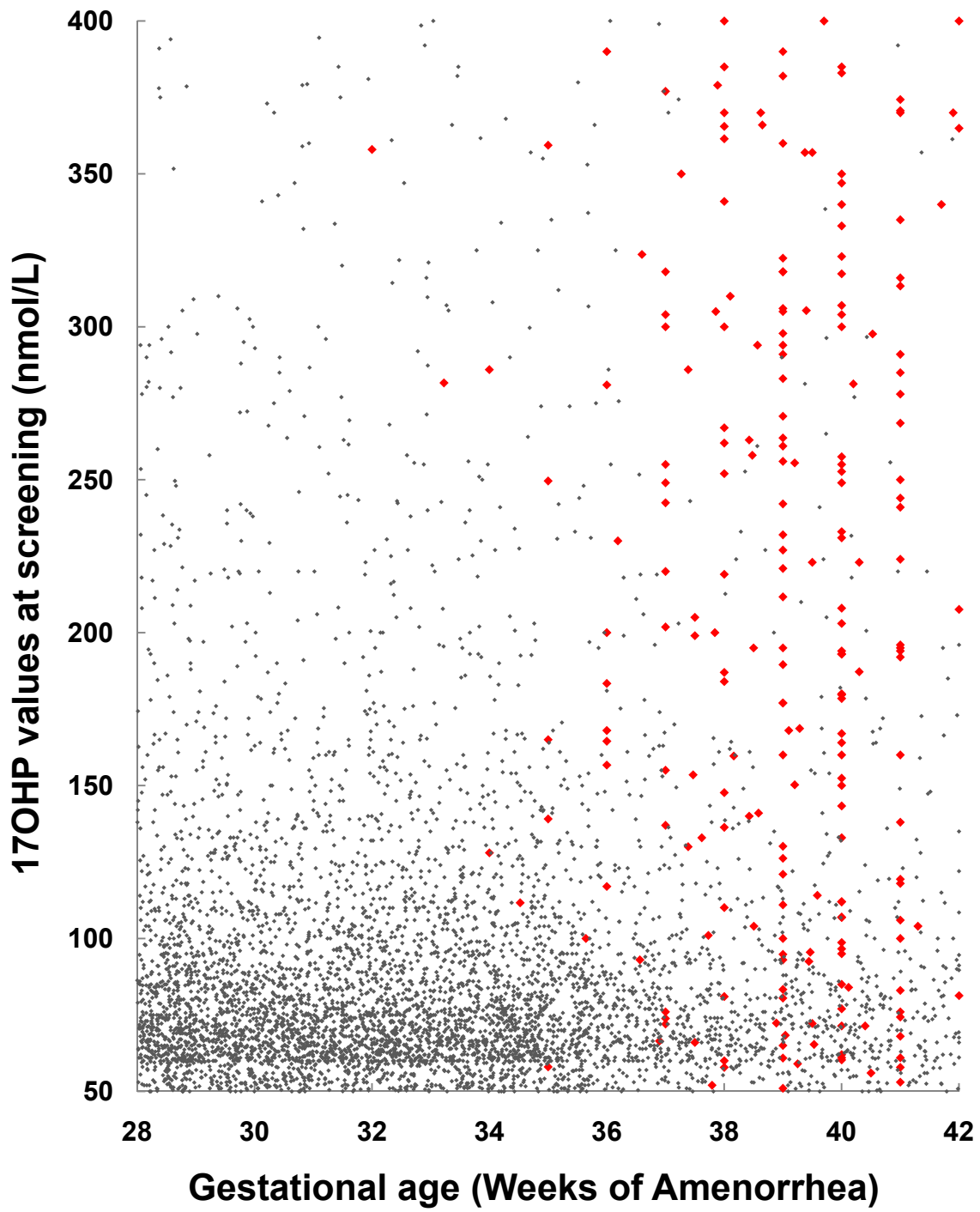


Figure 2

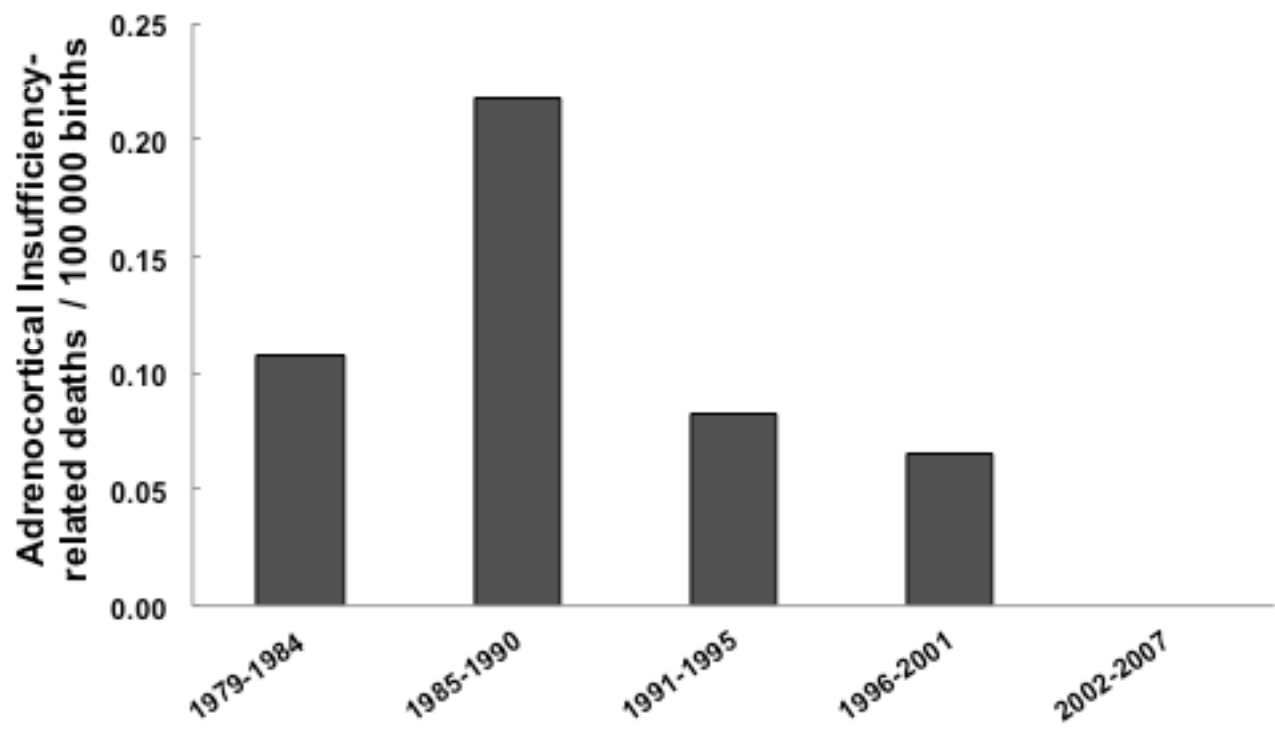


Figure 3