Supplementary data: Bayes' calculation for determining the probability of parent status for nonclassic congenital adrenal hyperplasia (NCCAH)

The probability of being affected with NCCAH ($p\{CC\}$) in the general population of unknown carrier status is estimated to be 1/1000 or 0.001.

In this calculation, it was assumed that alleles act independently; we denote the probability of an allele being unaffected as $p\{c\} = p$ and its probability of being affected as $p\{C\} = 1 - p = q$. Thus, for both alleles, $p\{cc\} = p^2$; $p\{Cc \text{ or } cC\} = 2pq$; and $p\{CC\} = q^2$. Collectively these three outcomes are clearly exhaustive, as can be checked by their cumulative probability: $p^2 + 2pq + q^2$, which simplifies to $(p + q)^2 = 1^2 = 1$.

Since $p\{CC\} = q^2 = 0.001$, $q = 0.0316$, or 3.2%, indicates the chances an allele is affected; the complement $p = 1 - q = 0.968 = 96.8\%$ is the probability of an unaffected allele. Hence the chances of being:

-unaffected on both alleles: $p\{cc\} = p^2 = 0.968^2 = 0.937$ or 93.7%;

-a carrier: $p\{Cc\} + p\{cC\} = 2pq = 2 * 0.968 * 0.032 = 0.062$ or 6.2% (where * indicates multiplication)

-being affected (i.e., being affected on both alleles) = $p\{CC\} = q^2 = 0.032^2 = 0.001$ or 0.1%.

Thus, the ratio of carriers to affected individuals to is $2pq / q^2 = 2p / q = 2\ (0.968) / 0.032 = 60.5$, indicating that for every 61 carriers, there is one affected individual. In a population of 1,000 people this would translate to one affected, 61 carriers, and 938 unaffected.

Let **Z** = a child in the study whose outcome is CC (there were no unaffected children in this study);
let **Y** be the child's genotyped parent (potentially with alleles: cc, Cc or cC, or CC);
let **X** be the other parent (potentially with alleles: cc, Cc or cC, or CC).
Then the following table shows the conditional probabilities that **Z** = CC given each possible combination of **X** and **Y**, assuming the parents' genetic contributions are independent:

		Y (genotyped)		
		(cc)	(cC or Cc)	(CC)
X	(cc)	***0***	***0***	***0***
	(cC or Cc)	***0***	***0.25***	***0.5***
	(CC)	***0***	***0.5***	***1***

From this table and the probabilities noted above for pairs of alleles, the conditional probabilities of the child being affected (**Z** = CC) given the various parental statuses of the genotyped parent **Y** are:

p(**Z** = CC | **Y** = cc): sum over column with **Y** = cc: $0 * p$(**X** =cc) + $0 * p$(**X** =cC or Cc) + $0 * p$(**X** = CC) = 0;

p(**Z** = CC | **Y** = cC or Cc): sum over column with **Y** = cC or Cc:
$= 0 * p$(**X** = cc) + $0.25 * p$(**X** =cC or Cc) + $0.5 * p$(**X** = CC) = $(0 * p^2) + (0.25 * 2pq) + (0.5 * q^2)$
$= (pq / 2) + (q^2 / 2) = q / 2(p+q) = q / 2 * 1 = q / 2$;

p(**Z** = CC | **Y** = CC): sum over column with **Y** = CC: $0 * p$(**X** =cc) + $0.5 * p$(**X** = cC or Cc) + $1 * p$(**X** =CC)
$= (0.5 * 2pq) + (1 * q^2) = pq + q^2 = q(p + q) = q * 1 = q$.

According to Bayes' Theorem, the conditional probability of the parent given an affected child is, therefore:
p(**Y** = CC | **Z** = CC) = [p(**Z** = CC | **Y** = CC) p(**Y** = CC)] / [p(**Z** = CC)] (a)

Assuming the prevalence in the general population, estimated to be 0.001, applies to both parents and children, we have: p(**Y** = CC) = p(**Z** = CC) = 0.001; thus equation (a) above reduces to:
p(**Y** = CC | **Z** = CC) = [p(**Z** = CC | **Y** = CC) * 0.001] / [0.001] = p(**Z** = CC | **Y** = CC) = q
In summary: the probability of a parent being affected, given an affected child, is q = 3.2%.

The probability of a parent being a carrier, given an affected child is (again by Bayes' theorem):

$p(\mathbf{Y} = \text{Cc or cC} \mid \mathbf{Z} = \text{CC}) = [p(\mathbf{Z} = \text{CC} \mid \text{Y} = \text{Cc or cC})\ p(\mathbf{Y} = \text{Cc or cC})] / [p(\mathbf{Z} = \text{CC})] = [\ (q / 2)\ (2pq)] / [q^2]$
$= (q / 2)\ (2pq / q^2) = 2pq^2 / 2q^2 = p = 0.968$, or 96.8%.

This calculation can be confirmed by simply noting that if the child is affected, the parent must be either affected or a carrier, so the probability of being a carrier is simply $1 - p$ (affected given child is affected) $= 1 - q$ (from above) $= 1 - 0.032 = 0.968$.

In the current study, we observed 10 / 249 (4.0%) parents were affected. Based on the calculations above, we expected 3.2% of the parents to be affected, given an affected child, and expected 96.8% of the parents to be carriers, given an affected child.

Please note that the calculations presented here do not consider multiple affected children per parent combination or other factors (e.g., ethnicity) that could alter the prior probabilities.