Supplementary Table SV **Assessment of study quality against Newcastle-Ottawa criteria for cohort studies; explanation of categorization is presented in supplementary material alongside its corresponding number.**

Author	Year	Representativeness of exposed cohort[a]	Selection of non-exposed cohort	Ascertainment of exposure	Demonstrated outcome was not present at start of study	Comparability of exposed and non-exposed	Assessment of outcome	Sufficient follow-up length[b]	Loss to follow-up
Jensen *et al.*	2010	Somewhat representative of mothers and boys in the community	Drawn from the same community as the exposed cohort	Written self-report and structured interview	Yes	Cohorts comparable (study controls for child's age and other factors)	Record linkage	Yes	2% lost to follow-up, unlikely to cause bias
Kristensen *et al.*	2011	Somewhat representative of mothers and boys in the community	Drawn from the same community as the exposed cohort	Written self-report or structured interview	Yes	Cohorts comparable (study controls for child's age and other factors)	Independent assessment	Yes	No statement[c]
Philippat *et al.*	2011	Somewhat representative of mothers and boys in the community	Drawn from the same community as the exposed cohort	Structured interview	Yes	Cohorts comparable (study controls for child's age and other factors)	Independent assessment	Yes	No statement
Rebordosa *et al.*	2008	Somewhat representative of mothers and boys in the community[d]	Drawn from the same community as the exposed cohort	Structured interview and Written self-report	Yes	Cohorts comparable (study controls for child's age and other factors)	Record linkage	Yes	1.48% lost to follow-up, unlikely to cause bias
Snijder *et al.*	2012	Somewhat representative of mothers and boys in the community	Drawn from the same community as the exposed cohort	Written self-report	Yes[d]	Cohorts comparable (study controls for child's age and other factors)	Independent assessment	Yes	6.9% lost to follow-up, unlikely to cause bias[e]

[a]This item is poorly defined by the authors of the Newcastle-Ottawa scale. For clarification, for the purposes of this review we have defined this criterion as whether the members of the cohort who were exposed to analgesia during pregnancy might (or might not) be considered representative of members of the general population who were exposed to analgesia during pregnancy. All study cohorts were deemed to be 'somewhat representative', although this is a relatively vague classification with little actual application. Overall, we did not consider that any of the included studies introduced meaningful bias by inadequately addressing this criterion.

[b]For the purposes of this review, follow-up time was deemed sufficient provided that cryptorchidism was defined at or sometime after birth (i.e. an inevitable characteristic of any congenital anomaly study).

[c]No information was provided regarding loss to follow-up between the point at which the questionnaire was administered (during the third trimester) and outcome measurement (the latter differing between participants within the study). The method of outcome measurement (but not loss to follow-up) for this cohort study is further described in by the authors in the earlier manuscript by Boisen *et al.* (2011): 'The boys were examined shortly after birth and at 3 months of age. In Denmark, all boys were re-examined at 18 months of age. In Finland, however, only boys with congenital cryptorchidism and selected controls were seen at 18 months. In case of preterm birth, examination took place at birth (Finland) or around the expected date of delivery (Finland and Denmark), or both (Finland), and again 3 and 18 months later.'

[d]While 898 mothers were enroled into this cohort study (the 'Generation R' study) at the birth of their child, none of these participants were included in the final analysis (since data were not collected during pregnancy for these mothers).

[e]The authors state that a total of 3982 boys were available for follow-up, of which 3708 boys visited a clinic for follow-up at least once (93.1%) and were thus included in the final cohort.