The incidence of childhood and adolescent seizures in the UK from 1999 to 2011; a retrospective cohort study using the Clinical Practice Research Datalink

Cormac J Sammon, Rachel A Charlton, Julia Snowball, John G Weil

1. Cormac J Sammon (Corresponding author)
   University of Bath,
   Claverton Down,
   Bath, BA2 7AY,
   UK
   C.Sammon@bath.ac.uk

2. Rachel A Charlton
   University of Bath,
   Claverton Down,
   Bath, BA2 7AY,
   UK
   R.A.Charlton@bath.ac.uk

3. Julia Snowball
   University of Bath,
   Claverton Down,
   Bath, BA2 7AY,
   UK
   J.Snowball@bath.ac.uk

4. John G Weil
   Novartis Vaccines and Diagnostics,
   Hullenbergweg 83
   Amsterdam,
   Netherlands
   john.weil@novartis.com

Keywords: Vaccines, Seizures, General Practice, Incidence, Epidemiology

Word count:

Abstract 265/300
Full text: 2738/3000
Tables/Figures: 6/6
Background

In postmarketing vaccine surveillance, adverse events observed in a vaccinated population are compared to the number expected based on a background incidence rate. The background rate should be accurate and obtained from a population comparable to the one vaccinated. Such rates are often not available.

Methods

The incidence rate of generalised convulsive, febrile and afebrile seizures was estimated in individuals born after 01-January-1998 and aged between 2 months and 15 years of age using the UK Clinical Practice Research Datalink (1999-2011).

Results

The study population consisted of 1,532,992 individuals (4,917,369 person years (PY) of follow up). A total of 28,917 generalised convulsive seizure events were identified during follow-up, the overall incidence rate was 5.88 per 1,000PY. Age specific rates increased sharply from 4/1,000PY at 2 months of age, peaked at 19/1,000PY at 16 months and decreased until approximately 6 years of age at which point they became relatively stable at 2/1,000PY. 67% of GCSs were categorised as febrile: 56% using Read codes, 11% using free text. Febrile seizures accounted for the age trend in GCS, with rates peaking at 16.1/1,000PY at 16 months of age while afebrile seizure rates remained relatively stable across all ages (24 seizures per 1,000PY). Analysis by first occurrence of febrile seizure showed a similar pattern, comparable to published studies on the incidence of seizures in childhood.

Discussion

The rates reported in this study could be used in the postmarketing surveillance of infant vaccines. However, given the variation across strata, and the potential underascertainment of seizure events presenting to A&E, care must be taken when interpreting and using these rates.
Generalised convulsive seizures (GCS) are episodes of neuronal hyperactivity resulting in involuntary muscle contraction and impairment or loss of consciousness. [1]. GCS can result from known causes such as epilepsy syndromes, central nervous system infections and acute electrolyte imbalances, however most remain idiopathic [2].

Febrile seizures (FS) are a particular type of idiopathic GCS in which seizure onset is preceded by a fever [3]. They are the most common type of seizure with up to 5% of the population suffering at least one FS before 5 years of age [4-6]. The incidence of febrile seizures is highest between 6 and 36 months of age and has been shown to peak at approximately 18 months of age [7]. Approximately one third of individuals who have a febrile seizure will suffer a recurrence, with family history, lower peak temperature of fever, short duration of fever and febrile seizures in young individuals being associated with a greater risk of recurrence [8, 9].

A number of vaccines have been linked with an increased risk of febrile seizure, most notably the measles, mumps and rubella (MMR) vaccine [10], the whole cell pertussis (DTP) vaccine [11] and the CSL pandemic (H1N1) influenza vaccine [12]. As a result, (febrile) seizures are commonly prioritised for close surveillance in vaccine safety monitoring.

Prior to the 2009 pandemic vaccination campaign Black et al highlighted the importance of background incidence rates in vaccine surveillance [13]. They illustrated how such rates can be used as an expected rate against which the rates observed in a vaccinated population could be compared. They also underlined the need to be aware of any geographic, ethnic and age differences in such rates and their dependence on the method used to develop the rates. Examples of their use during the 2009 pandemic can be found in the literature [14, 15]. Background incidence rates of GCS for the United Kingdom are not widely available in the published literature [4, 16] and they lack the appropriate level of stratification for use in an infant vaccine programme. This study seeks to investigate the incidence of GCS in infants, children and adolescents in the UK Clinical Practice Research Datalink (CPRD) and to determine whether the CPRD can accurately categorise GCS by distinguishing FS from afebrile GCS.
Methods

This study was carried out using the UK Clinical Practice Research Datalink (CPRD). The CPRD is an electronic healthcare database containing the anonymised primary care medical records of ~8.4% of the UK general population. Patient data routinely collected in primary care and therefore available in the database include demographic details, diagnoses and symptoms including those leading to hospital admissions, immunisations, pregnancies, laboratory tests, referrals to specialists, prescriptions issued by the general practitioner (GP, primary care physician) and deaths [17, 18]. In the UK a patient's primary care record is considered their main electronic health record therefore events occurring in secondary care (e.g. emergency room visits, hospital inpatient events) should be reported to a patient’s GP and entered into their record. Despite this, recording of secondary care data in the primary care record is incomplete. Clinical events in the CPRD are recorded using clinical codes known as a Read codes. There are currently over 100,000 Read codes each of which is associated with a short description of varying specificity. Diagnostic Read codes can be considered equivalent to ICD codes, with many mapping directly to specific ICD codes, however a range of additional Read codes exist to facilitate the complexities of patient management in primary care. In order to further facilitate the management of patients in primary care, recording of additional, unstructured textual information in association with a Read code is also possible. This information, commonly referred to as ‘free text’, generally contains elaborations on the information in the coded record.

The study period ran from 01-January-1999 to the 31-December-2011. The study population comprised individuals permanently registered in the CPRD and aged between 2 and 180 months (15 years) at some point during the study period. Follow up of each patient began at the start of the study period, an age of 2 months or the date of registration with the CPRD, whichever was latest. Follow up ended at the end of the study period, an age of 180 months, date of death or transfer out of the GP practice, whichever is earliest.
In the CPRD the month of birth is only available for individuals aged less than 15 years old. Individuals born before the 01-January-1998 had reached the age of 15 before the end of the study period and therefore did not have a record of their month of birth available. Individuals were therefore excluded from the study population if they were born before 01-January-1998.

Our definition of GCS generally followed the Brighton Collaboration definition [1] in that it sought to include all convulsive seizures regardless of their cause and nature. Primary care data, such as that contained in the CPRD, is often unsuited to classification of cases at specific Brighton collaboration classification levels as data on many of the necessary classification criteria are not commonly recorded in general practice. As a result we did not attempt to define cases at specific Brighton collaboration classification levels. Operationally, GCS events were therefore identified as any event recorded against one of the seizure related Read codes listed in tables A1.1, A1.2 or A1.3 in supplementary file 1.

We sought to separate GCS events into those that were febrile and afebrile. Two clinical definitions of febrile seizure can be found in the published literature [3, 19]. Both of these define a febrile seizure as a seizure which is associated with a fever and occurs in an individual aged less than 5 years old with no central nervous system infections and with no history of afebrile seizure. The primary difference between these two definitions is the minimum age at which they determine a seizure can be defined as febrile (1 vs. 3 months). Taking both the nature of CPRD data and these clinical definitions of febrile seizures into account, in this study a febrile seizure was defined as any seizure occurring in association with a fever in an individual: aged greater than 1 month old, with no evidence of central nervous system infection and with no history of epilepsy. Note that the exclusion of all individuals with a history of afebrile seizure was not included in this definition as the sensitivity of such an exclusion criterion in the CPRD is likely to be poor. Operationally, febrile seizures meeting this definition were identified in the CPRD using (a) a code for febrile seizure, (b) a code for seizure and a code for fever or febrile seizure recorded within 2 weeks either side (c) a code for seizure and a free text entry indicating a fever was present. In line with our clinical definition of febrile seizure, events identified under
definitions (a), (b) and (c) were not considered febrile if central nervous system infections were
recorded in the patient’s record in the 2 weeks before or 6 weeks after the event or if an epilepsy code
was recorded anywhere prior to the event. The codes defining such events are provided in Table A1.5
and Table A1.6 of supplementary file 1.

All GCS events that did not meet the above definition of a febrile seizure were considered afebrile
seizure events.

The total number of GCS events in age, sex and calendar year specific periods were calculated and
used as the numerators in the stratified incidence rates. The amount of person-time contributed by
the study population in each age, sex and calendar year specific period was calculated and used as the
denominator in the stratified incidence rates. Unless otherwise specified, all incidence rates were
reported as numbers of seizures per 1,000 person years (PY). Confidence intervals were estimated
assuming a normal binomial distribution for all rates. As per CPRD policy, strata with <5 events are
reported as <5 and no incidence rates were calculated. Rates of afebrile and febrile seizures were
calculated using the same method. In addition, rates of first GCS, febrile and afebrile seizure were
estimated by including only the first event per individual in the numerator and censoring follow-up for
the denominator at first seizure occurrence. The total number of events and the number of first events
that could be expected in given time periods after vaccination of a hypothetical cohort of children was
also calculated by multiplying the incidence rates by the parameters describing the hypothetical
population.

Results

The study population consisted of 1,173,916 individuals who contributed 4,917,369 person years of
follow up during the study period. A total of 28,917 GCS events were identified during follow-up
providing an overall incidence rate of GCS in the entire study population of 5.88 per 1,000 PY. Figure
1 and Table 1 show age specific incidence rates of GCS. Rates increased sharply from 3.5/1,000 PY at
2 months of age, peaked at 19.2/1,000 PY at 16 months and decreased until approximately 6 years of age at which point they became relatively stable at approximately 2/1,000 PY. The incidence rate for those aged between 2 months and 5 years was 8.99 per 1,000 PY.

No meaningful differences were observed between the sex specific rates of GCS (data not shown). Table A2.1 (Supplementary file 2) shows the age category specific incidence of GCS across calendar years. The distribution by age was similar within calendar years. However, rates decreased slightly over time.

There were 19,622 febrile seizure events and 9,295 afebrile seizure events resulting in overall incidence rates of 4.01 and 1.89 seizures per 1,000 PY respectively. Figure 1 and Table 1 compare age specific incidence rates of febrile, afebrile and generalised convulsive seizure. The rate of afebrile seizure remained stable across ages (2-4 seizures per 1,000 person years). In contrast, febrile seizure rates account for the peak in the GCS incidence rate at 18 months and the general trend observed up to 6 years. After 5 years of age the rate of febrile seizures is lower than that of afebrile. The incidence rate for those aged between 2 months and 5 years was 6.68 per 1,000 PY for febrile seizure and 2.32 per 1,000 PY for afebrile seizure. No meaningful sex distribution was observed for febrile/afebrile seizure (data not shown). Table 3 describes the proportion of febrile seizures that were identified using each of the three case identification methods.

When the analysis was restricted to first events 18,336 GCS events were identified during 4,837,363 person years of follow up, 14,015 febrile seizure events were identified during 4,839,251 person years of follow up and 5,447 afebrile seizure events were identified during 4,895,611 person years of follow up. This resulted in overall incidence rates of 3.79, 2.90 and 1.11 seizures per 1,000 PY respectively. Figure 2 and Table 2 show age specific incidence rates of first febrile, afebrile and generalised convulsive seizure. While the age specific rates of first seizure are unsurprisingly lower than those for all seizures, the distribution across age categories is similar to that for all seizures. Gender and year
specific rates of first seizure showed similar distributions to rates for all events (Table A2.2, supplementary file 2).

Supplementary file 3 contains interactive tables which can be used to obtain the incidence rates and expected numbers of first/all febrile, afebrile or generalised convulsive seizure events for reader-defined age, sex and calendar year stratum, in addition the reader can alter the number of individuals vaccinated and the duration of surveillance.

Discussion

This study reports month of age specific incidence rates of generalised convulsive, febrile and afebrile seizures for children and adolescents in the CPRD, the incidence rates can be used to calculate the expected numbers of seizures in an infant childhood vaccination programme.

The analysis by first occurrence of GCS was included to compare the results for the categorisation of febrile and afebrile seizures with the literature [5, 20]. The pattern is similar with FS less common than afebrile seizures in the first few months of life, increasing thereafter to reach a peak in the second year of life, then dropping sharply from the third year of life to become less frequent than afebrile seizures after four or five years of age. The results suggest the CPRD can reliably identify FS. Our use of diagnostic codes associated with fever identified 4% of FS, while our use of free text strings associated with fever identified 16% of FS (Table 3). While we have not validated the febrile nature of seizures identified using each of the methods we believe that similar approaches should be considered when seeking to distinguish febrile seizures from afebrile seizures on the CPRD. If specificity were preferable, the case definition used could be amended to include only a record of fever within seven days of the record of seizure.

The CPRD provides a large, well-defined source population which has been shown to be representative of the age and sex distribution of the UK population [21, 22]. A disadvantage of the CPRD is that
seizures may not always be recorded in a patient’s general practice record using a relevant Read code when patients with seizures present directly to Accident and Emergency (A&E) services (secondary care). Comparing CPRD incidence rates to those reported in the literature (Table 4) the incidence rates we observed in the CPRD are between 10% and 88% lower than those observed in prospective follow-up studies [20, 23, 24] and electronic health record studies in the Danish National Hospital Register [11, 16] that all included A&E data. The exception is a third Danish study, which also used the Danish National Hospital Register (but only included primary discharge diagnoses) [25]. This illustrates the importance of understanding the methods used to produce the rates. The impact of A&E events is most evident in the VAESCO [16] data that used standardized methodology across several European electronic health record databases and shows a markedly higher incidence in the Danish data source that included A&E data.

In future, ascertainment of seizure events on the CPRD might be improved through linkage with Hospital Episode Statistics (HES) data; however, the quality of HES A&E data remains questionable [26]. In the absence of such linkage, an alternative option to increase the ascertainment of seizures presenting in secondary care would be to search the free text associated with Read codes such as “discharge summary”, “hospital discharge letter” or “Seen in A&E” for key words related to a seizure. This would allow for the identification of events in which the general practice has been informed of event occurrence but has recorded the event in the free text associated with a non-seizure diagnostic code rather than using a seizure-specific code.

False positive outcome misclassification may also have occurred in this study, for example if diagnostic codes for seizure had been used to record follow-up consultations about an earlier recorded seizure episode. While we have not formally investigated the extent of positive or negative misclassification, overall, we expect the magnitude of false negative misclassification to outweigh any false positive misclassification and therefore suggest that our rates be assumed to under-estimate the true rate of seizures in the population. Additionally, the event dates used in this study may reflect the date an
event was recorded in the database rather than the date of event occurrence; this may have a small
impact on our age and calendar year specific rates.

In observed vs expected calculations the number of subjects exposed to the vaccine is often the least
accurate of the inputs, and similarly the results are generally more sensitive to the risk window
selected than to variations in the background rate. Nevertheless it is important to select the most
accurate background rate available. If the required age strata are available from a data source in a
different geographic region which includes A&E events then using these as inputs would be preferable
to using a geographically accurate source which lacks A&E events as Table 3 shows less variation due
to geography (within Europe) than to inclusion of A&E events.

However, given the variation observed in incidence rates across age strata it is important that
appropriate consideration be given to which is the most comparable age stratum to use. However, the
specific age stratum of interest is often unlikely to be available in the literature. The interactive
spreadsheet we provide in Supplementary File 3 should allow the reader to obtain incidence rates and
expected numbers of events for the stratum that is most comparable to the one from which their
observed events are derived; it also allows sensitivity analyses assuming a level of underestimation,
whereby the reader can specify a percentage of seizures they assume the CPRD has missed when
calculating expected numbers of events. Such a resource should prove useful in future studies
reporting background incidence rates for use in post-marketing surveillance.

Conclusions

The results reported in this study provide reference or background incidence rates that can be used
in the postmarketing surveillance of vaccines and other products with the potential to cause seizures.
However, given the potential for underascertainment of seizure events, and the variation observed
across age strata, care must be taken when interpreting and using these rates.
Authorship

CJS, RAC and JGW designed the study, CJS and JS extracted and analysed the data, CJS drafted the manuscript, RAC, JGW and JS revised the manuscript for important intellectual content. CJS, RAC, JS and JGW read and approved the final manuscript submitted for publication.

Acknowledgements

This study is based in part on data from the Full Feature General Practice Research Database (now termed Clinical Practice Research Database) obtained under licence from the UK Medicines and Healthcare Products Regulatory Agency. However, the interpretation and conclusions contained in this study are those of the authors alone. The authors would like to acknowledge the valuable contribution of Professor Corinne de Vries in the early stages of this study.

Conflict of interest

This study was funded by Novartis Vaccines and Diagnostics S.r.l., study number V72_630BTP. The author affiliated with Novartis Vaccines and Diagnostics S.r.l. was involved in study design and preparation of the manuscript. However, this study does not evaluate the safety of any products.

Ethical Statement

The GPRD has a single Multi-Centre Ethics approval for all observational studies using GPRD data (Trent MREC, ref: 05/MRE04/87). The extraction and analysis of the data used in this study has been approved by the Independent Scientific Advisory Committee of the CPRD under approval number 10_057.


Table 1. Age category specific rates of febrile seizure, afebrile seizure and generalised convulsive seizure (GCS). Includes all events occurring in the study population during follow up (i.e. multiple events per individual included).

<table>
<thead>
<tr>
<th>Age category (months)</th>
<th>GCS</th>
<th>Febrile</th>
<th>Afebrile</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>IR</td>
<td>Cl95</td>
</tr>
<tr>
<td>2-12</td>
<td>4768</td>
<td>8.34</td>
<td>(8.11-8.58)</td>
</tr>
<tr>
<td>13-24</td>
<td>10341</td>
<td>16.63</td>
<td>(16.31-16.95)</td>
</tr>
<tr>
<td>25-60</td>
<td>9933</td>
<td>6.13</td>
<td>(6.01-6.25)</td>
</tr>
<tr>
<td>61-120</td>
<td>3290</td>
<td>1.95</td>
<td>(1.88-2.02)</td>
</tr>
<tr>
<td>121-180</td>
<td>585</td>
<td>1.41</td>
<td>(1.29-1.52)</td>
</tr>
</tbody>
</table>

Table 2. Age category specific rates of first febrile seizure, afebrile seizure and generalised convulsive seizure (GCS). Includes only the first event per individual during follow up (i.e. multiple events per individual not included).

<table>
<thead>
<tr>
<th>Age category (months)</th>
<th>GCS</th>
<th>Febrile</th>
<th>Afebrile</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>IR</td>
<td>Cl95</td>
</tr>
<tr>
<td>2-12</td>
<td>3659</td>
<td>6.42</td>
<td>(6.21-6.63)</td>
</tr>
<tr>
<td>13-24</td>
<td>7232</td>
<td>11.76</td>
<td>(11.49-12.03)</td>
</tr>
<tr>
<td>25-60</td>
<td>5502</td>
<td>3.46</td>
<td>(3.37-3.55)</td>
</tr>
<tr>
<td>61-120</td>
<td>1637</td>
<td>0.99</td>
<td>(0.94-1.04)</td>
</tr>
<tr>
<td>121-180</td>
<td>306</td>
<td>0.75</td>
<td>(0.67-0.84)</td>
</tr>
</tbody>
</table>
Table 3 Proportion of febrile seizures identified using different case identification methods

<table>
<thead>
<tr>
<th>Identification method</th>
<th>N cases</th>
<th>% total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Read code for Febrile Seizure</td>
<td>15,733</td>
<td>(80.1)</td>
</tr>
<tr>
<td>Read code for seizure and Read code for fever or febrile seizure code within 7 days</td>
<td>600</td>
<td>(3.1)</td>
</tr>
<tr>
<td>Read code for seizure and Read code for fever or febrile seizure code within 14 days</td>
<td>152</td>
<td>(0.8)</td>
</tr>
<tr>
<td>Read code for seizure and free text containing a string related to fever* within 7 days</td>
<td>2,830</td>
<td>(14.4)</td>
</tr>
<tr>
<td>Read code for seizure and free text containing a string related to fever* within 14 days</td>
<td>307</td>
<td>(1.6)</td>
</tr>
<tr>
<td>Febrile seizures</td>
<td>19622</td>
<td>(100)</td>
</tr>
</tbody>
</table>

*Text strings related to fever include “*fever*”, “*febrile*”, “*pyrex*”, “*temp*”

Table 4. Comparison with the literature
<table>
<thead>
<tr>
<th>Study</th>
<th>Period children born in</th>
<th>Country</th>
<th>A&amp;E</th>
<th>First/All events</th>
<th>Study perspective</th>
<th>Age (months old)</th>
<th>Outcome</th>
<th>IR (per 1000PY)</th>
<th>Equivalent CPRD IR</th>
</tr>
</thead>
<tbody>
<tr>
<td>VAESCO [16]</td>
<td>1993-&lt;2009</td>
<td>UK</td>
<td>No</td>
<td>All[^d]</td>
<td>Retrospective</td>
<td>0-59</td>
<td>GCS</td>
<td>7.4</td>
<td>9.0</td>
</tr>
<tr>
<td>VAESCO [16]</td>
<td>1993-&lt;2009</td>
<td>Spain</td>
<td>No</td>
<td>All[^d]</td>
<td>Retrospective</td>
<td>0-59</td>
<td>GCS</td>
<td>5.0</td>
<td>9.0</td>
</tr>
<tr>
<td>VAESCO [16]</td>
<td>1999-&lt;2008</td>
<td>Finland</td>
<td>No</td>
<td>All[^d]</td>
<td>Retrospective</td>
<td>0-59</td>
<td>GCS</td>
<td>8.9</td>
<td>9.0</td>
</tr>
<tr>
<td>Rasmussen et al [25]</td>
<td>1980-&lt;2010</td>
<td>Denmark</td>
<td>Yes</td>
<td>First</td>
<td>Retrospective</td>
<td>12-47</td>
<td>FS</td>
<td>5.9</td>
<td>6.1</td>
</tr>
<tr>
<td>Van den Berg et al [20]</td>
<td>1960-&lt;1968</td>
<td>USA</td>
<td>Yes</td>
<td>First</td>
<td>Prospective</td>
<td>0-59</td>
<td>GCS</td>
<td>6.5</td>
<td>6.0</td>
</tr>
<tr>
<td>Annegers et al [5]</td>
<td>1935-&lt;1985</td>
<td>USA</td>
<td>Yes</td>
<td>First</td>
<td>Retrospective</td>
<td>0-59</td>
<td>FS</td>
<td>2.0%^[^g]</td>
<td>2.2%</td>
</tr>
<tr>
<td>Verburgh et al [24]</td>
<td>1982-&lt;1988</td>
<td>Netherlands</td>
<td>No</td>
<td>First</td>
<td>Prospective</td>
<td>2-59</td>
<td>FS</td>
<td>5.5</td>
<td>4.9</td>
</tr>
</tbody>
</table>

[^a] A&E = Accident and Emergency  
[^b] GCS = Generalised convulsive seizures, FS = Febrile seizures, AFS = afebrile seizure  
[^c] Rates observed for similar outcome/age band in current study, however where rates in literature included events occurring before 2 months of age the CPRD rate will not.  
[^d] In the VAESCO study recurrent events within 2 weeks of a first event were excluded.  
[^e] Data collected through parental interview therefore A&E events assumed to be included.  
[^f] Incidence rate not available for these studies, prevalence shown instead, and CPRD prevalence for comparison is calculated among subset of individuals with complete follow up from 2 months up to 5 years of age (n= 237,832, ~20% of total study population).  
[^g] Verity et al included children born in a single week in April 1970.  
[^h] Only primary discharge diagnoses included.
**Figure 1** Incidence of febrile, afebrile and generalised convulsive seizures by month of age.

**Figure 2** Incidence of first febrile, afebrile and generalised convulsive seizures by month of age.