A systematic approach for identifying shared mechanisms in epilepsy and its comorbidities

Charles Tapley Hoyt1,2,†,*, Daniel Domingo-Fernández1,2,†, Nora Balzer2, Anka Güldenpfennig1, and Martin Hofmann-Apitius1,2

1. Department of Bioinformatics, Fraunhofer Institute for Algorithms and Scientific Computing, Sankt Augustin 53754, Germany
2. Bonn-Aachen International Center for IT, Rheinische Friedrich-Wilhelms-Universität Bonn, Bonn 53113, Germany

*Corresponding Author: Hoyt, C. T. Department of Bioinformatics, Fraunhofer Institute for Algorithms and Scientific Computing, Sankt Augustin 53754, Germany. Telephone details; +49 2241 14-2268
† These authors contributed equally to this work.

Keywords: epilepsy; Alzheimer's disease; comorbidity; systems biology; pleiotropy; neurodegeneration; epidemiology

Abstract

Cross-sectional epidemiological studies have shown that the incidence of several nervous system diseases is more frequent in epilepsy patients than in the general population. Some comorbidities (e.g., Alzheimer's disease and Parkinson's disease) are also risk factors for the development of seizures; suggesting they may share pathophysiological mechanisms with epilepsy.

A literature-based approach was used to identify gene overlap between epilepsy and its comorbidities as a proxy for a shared genetic basis for disease, or genetic pleiotropy, as a first effort to identify shared mechanisms. While the results identified neurological disorders as the group of diseases with the highest gene overlap, this analysis was insufficient for identifying putative common mechanisms shared across epilepsy and its comorbidities. This motivated the use of a dedicated literature mining and knowledge assembly approach in which a cause-and-effect model of epilepsy was captured with Biological Expression Language.

After enriching the knowledge assembly with information surrounding epilepsy, its risk factors, its comorbidities, and anti-epileptic drugs, a novel comparative mechanism enrichment approach was used to propose several downstream effectors (including the GABA receptor, GABAergic pathways, etc.) that could explain the therapeutic effects of carbamazepine in both the contexts of epilepsy and AD.

We have made the Epilepsy Knowledge Assembly available at https://www.scai.fraunhofer.de/content/dam/scai/de/downloads/bioinformatik/epilepsy.bel and queryable through NeuroMMSig at http://neurommsig.scai.fraunhofer.de. The source code used for analysis and tutorials for reproduction are available on GitHub at https://github.com/cthoyt/epicom.
Introduction

Seizures are transient occurrences of signs and symptoms due to abnormal or excessive neuronal activity in the brain [1]. Classically, their underlying causes were thought to be the primary drivers for increasing mortality in epilepsy. While epilepsy has been classically studied as a disorder of the brain characterized by an enduring predisposition to epileptic seizures, it is no longer considered a condition in which seizures are the only concern [2]. Epilepsy is also associated with several comorbidities, including Alzheimer's disease (AD), Parkinson's disease (PD), other nervous system diseases, and psychiatric disorders [3-5], due to a variety of genetic, biological, and environmental factors [6].

The prevalence of migraine in epilepsy patients (under 64 years old) is 5.71% in contrast to 3.47% in the general population [7]. The mechanistic understanding of epilepsy and migraine presumes that they share the underlying pathophysiology related to alterations in sodium and calcium ion channels and ion transporters (sodium-potassium transport) [8,9]. Moreover, drugs acting on voltage-gated sodium channels and γ-aminobutyric acid (GABA) receptors (e.g., valproate, topiramate, etc.) are used not only prevent migraine attacks, but are also used as antiepileptic drugs [10,11].

The prevalence of epileptic seizures in AD patients is strongly influenced by genetic factors [12] — epilepsy and seizures occur more often in patients with early-onset familial AD than those with sporadic AD [13]. Further, convulsive seizures have been described in approximately 40% of familial AD patients with the PSEN1 p.Glu280Ala mutation [14], 30% with PSEN2 mutations [15], and 57% with amyloid precursor protein (APP p.Thr174Ile) duplications [16]. Additionally, the p.Pro86Lys mutation in CALHM1 (rs11191692) is associated with both AD and temporal lobe epilepsy through its influence on calcium homeostasis [17].

Evidence that some comorbidities (e.g., AD and depression) also act as risk factors for developing seizures suggests they may share pathophysiological mechanisms [13, 18]. Conversely, epileptic seizures (as well as neoplasms) have been reported to cause intellectual disabilities in patients with tuberous sclerosis (Jansen et al., 2008; Osborne et al., 2010). Furthermore, certain antiepileptic drugs (e.g., topiramate) are associated with higher incidence of cognitive problems [1, 21].

The several causal and associative relationships observed between epilepsy and other indications on the phenotypic level warrant further investigation for shared elements on the genetic and molecular levels. While inquiry on the genetic level often begins with genome-wide association studies to identify shared loci, identifying the appropriate data set(s) and linking intergenic single nucleotide polymorphisms (SNP) to their functional consequences across scales in complex disease is still a significant challenge [22]. Even after identifying disease-associated genes, it is difficult to assess their individual contributions to the complex etiology of epilepsy and its related indications. Thus, capturing the different causal relationships between biological entities involved in the pathophysiology of a complex disease is an essential step to a better understanding of the processes that lead to the disease state.

Here, we present two methods to hypothesize shared mechanisms: first, a literature-based approach for quantifying gene overlap between epilepsy and its comorbidities as a proxy for a shared genetic basis of disease, or genetic pleiotropy; and second, a systematic approach using the NeuroMMSig mechanism enrichment server.
Finally, we use these methods to propose an explanation for the observed therapeutic effects of a drug that has been studied in the contexts of epilepsy and AD. A schematic representation of the methodology, analysis, and results is presented in Figure 1.

**Figure 1: A graphical abstract of the methodology, analysis, and results presented in this work.** The two upper boxes represent the methodology while the two lower boxes represent the analysis and results. The upper-left box outlines the quantification of gene overlap between epilepsy and its well-known comorbidities described in Keezer *et al.* [23] (e.g., Alzheimer's disease, Parkinson's disease, etc.) using literature based methods. The upper-right box outlines the assembly of knowledge from epilepsy literature with chemical and pathway enrichment as described in the "Preparation for Mechanism Enrichment" and "Relation Extraction" sections. The lower-right box represents the comparative mechanism enrichment that was used to generate comorbidity insights (lower-left box) after literature-based methods proved insufficient.

**Materials and Methods**

**Quantification of Gene Overlap**

SCAIView v1.7.3, indexed from MEDLINE on 2016-07-14, ([http://academia.scaiview.com/academia](http://academia.scaiview.com/academia)) was used to identify and quantify the overlap of genes co-occurring in the literature with epilepsy and its novel and well-known comorbidities reviewed by Keezer *et al.* [23]. At least one disease was selected for each of the ICD (International Statistical Classification of Diseases and Related Health Problems) groups that were analyzed by Keezer *et al.* [23] as well as the addition of Parkinson's disease due to its previously published relevance [24].

In order to later assess publication bias in literature-based methods, we report the total number of documents associated with each disease by querying SCAIView with their corresponding MeSH terms and the total number of associated genes that have a positive relative entropy (i.e., occur more frequently in the results of a given SCAIView query than in the rest of the SCAIView indexed literature) in the context of the query as described by Younesi *et al.* [25] Gene sets for each comorbidity were then retrieved by constructing queries...
using their corresponding MeSH terms joined with the epilepsy MeSH term by the “AND” operator. The associated genes for each comorbidity query were identified with the same method and the epilepsy pleiotropy rate was calculated as the percentage of the genes associated with the comorbidity query in the set of associated genes with epilepsy (Table 1).

For example, 226 genes were found to be associated with diabetes using the comorbidity query [MeSH Disease:"Epilepsy"] AND [MeSH Disease:"Diabetes Mellitus"]. Of these, 184 had positive relative entropies, which indicate that these genes occur more frequently in literature mentioning both diseases than in the rest of the SCAIVView indexed literature. Finally, the epilepsy pleiotropy rate was calculated normalizing 184 by the total number of genes (2901) with a positive relative entropy found by querying for epilepsy [MeSH Disease:"Epilepsy"], resulting in a pleiotropy rate of 6.34%.

<table>
<thead>
<tr>
<th>Disease (MeSH ID)</th>
<th>Associated documents</th>
<th>Disease-associated Genes</th>
<th>Comorbidity-associated genes</th>
<th>Epilepsy pleiotropy rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Epilepsy (D004827)</td>
<td>192245</td>
<td>2901</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Alzheimer's Disease (D000544)</td>
<td>109495</td>
<td>4968</td>
<td>396</td>
<td>13.65</td>
</tr>
<tr>
<td>Migraine (D008881)</td>
<td>30928</td>
<td>1230</td>
<td>306</td>
<td>10.54</td>
</tr>
<tr>
<td>Parkinson's Disease (D010300)</td>
<td>79103</td>
<td>3646</td>
<td>258</td>
<td>8.89</td>
</tr>
<tr>
<td>Hypertension (D006973)</td>
<td>391190</td>
<td>5574</td>
<td>252</td>
<td>8.68</td>
</tr>
<tr>
<td>Dementia (D003704)</td>
<td>183802</td>
<td>5833</td>
<td>220</td>
<td>7.58</td>
</tr>
<tr>
<td>Diabetes Mellitus (D003920)</td>
<td>394411</td>
<td>6661</td>
<td>184</td>
<td>6.34</td>
</tr>
<tr>
<td>Anxiety (D001007)</td>
<td>84138</td>
<td>1782</td>
<td>124</td>
<td>4.27</td>
</tr>
<tr>
<td>Arthritis (D001168)</td>
<td>259327</td>
<td>5367</td>
<td>122</td>
<td>4.20</td>
</tr>
<tr>
<td>Cataract (D002386)</td>
<td>52150</td>
<td>2238</td>
<td>119</td>
<td>4.10</td>
</tr>
<tr>
<td>Colonic Neoplasms (D003110)</td>
<td>107274</td>
<td>3646</td>
<td>30</td>
<td>1.03</td>
</tr>
<tr>
<td>Urinary Incontinence (D014549)</td>
<td>34170</td>
<td>720</td>
<td>24</td>
<td>0.82</td>
</tr>
<tr>
<td>Peptic Ulcer (D010437)</td>
<td>68234</td>
<td>1445</td>
<td>21</td>
<td>0.72</td>
</tr>
<tr>
<td>Pulmonary Disease, Chronic Obstructive (D020424)</td>
<td>35627</td>
<td>2244</td>
<td>15</td>
<td>0.51</td>
</tr>
</tbody>
</table>

Table 1: Results of the Epilepsy comorbidity analysis using SCAIVView. The number of associated documents (column 2) retrieved from SCAIVView for each disease are shown given a reference query using corresponding the MeSH term from column 1. The disease-associated genes (column 3) contains the number of genes relevant to the corpus retrieved from a disease-specific query. The comorbidity-associated genes (column 4) contains the number of genes relevant to the comorbidity query between the target disease and epilepsy. Lastly, the epilepsy pleiotropy rate (column 5) describes the ratio of the count of genes reported in column 4 with the total number of epilepsy-associated genes (2901).

Relation Extraction

While literature co-occurrence can generate initial hypotheses about genetic pleiotropy, it does not provide sufficient mechanistic insight to explain the clinical observations in epilepsy and its comorbidities. The increasing quantity of knowledge in the biomedical domain makes it difficult or impossible for researchers to be knowledgeable in any but incredibly specific topics [26]. The task of manually generating pleiotropy hypotheses that explain overlap between the aetiological mechanisms of epilepsy and its comorbidities is daunting. In order to enable computer-aided automatic reasoning, the knowledge surrounding epilepsy and its comorbidities was
systematically extracted from the literature using manual relation extraction and encoded in Biological Expression Language (BEL) [27].

First, a corpus was generated from the 192245 documents related to epilepsy retrieved (Table 1) to further investigate the causal relations surrounding the identified genes. A second corpus was generated from the 2666 documents retrieved by querying SCAIView for “epilepsy” and its sub-terms in the Epilepsy Ontology [28] occurring with the free text, “comorbidity.” Manual relation extraction and encoding in BEL was then performed starting with a select subset of the two corpora based on their prioritization by SCAIView to generate the Epilepsy Knowledge Assembly. The knowledge assembly was further enriched with pharmacological knowledge surrounding 19 antiepileptic drugs and their targets from the Pharmacogenetics and Pharmacogenomics Knowledge Base (PharmGKB) [29]. Ultimately, the knowledge assembly comprised relations from 641 unique citations. Finally, the PyBEL framework [30] was used to parse and validate the syntax and semantics of the underlying BEL Script. A summary of the contents of the Epilepsy Knowledge Assembly is presented in Table 2.

Preparation for Mechanism Enrichment

The Epilepsy Knowledge Assembly was enriched with mechanistic annotations following the procedures outlined by Domingo-Fernández et al. [24] in order to integrate it into NeuroMMSig and enable multi-modal mechanism enrichment analyses with queries over genes, SNPs, and neuroimaging features.

A taxonomy of epilepsy mechanisms was generated by combining the list of well-established epilepsy mechanisms from Staley [31] with concepts from the Pathway Terminology System [32] co-occurring in articles matching either [MeSH Disease:"Epilepsy"] or entries in the Epilepsy Ontology indexed by SCAIView. The resulting 784 terms representing mechanisms were curated in order to normalize entities, remove irrelevant entries, and group similar terms. Next, relations in the Epilepsy Knowledge Assembly were annotated with mechanisms based on whether their entities were involved in the mechanism as outlined by Domingo-Fernández et al. [24]. During the annotation process, new mechanisms not yet included in the inventory were found in the literature; thus, the mechanism inventory was parallely updated until concluding the annotation with a total of 32 annotated subgraphs (Table 2). More details and examples about the mapping procedure can be found on the NeuroMMSig introduction page.

In the next section, we use NeuroMMSig to identify shared mechanisms between epilepsy and AD on the basis of the mechanism of action of multi-indication drugs.

Results and Discussions

Investigation of Comorbidities

Of the 2901 genes identified as associated with epilepsy by having a positive relative entropy score, Table 1 shows that nervous system disorders had the highest gene overlap (10.54% co-occurred with migraine, 7.58% with dementia, 8.89% with PD, and 13.65% with AD). These rates do not positively correlate with previous epidemiological studies of comorbidity incidence with the same conditions reviewed by Keezer et al. [23], therefore, literature-based gene overlap is a poor proxy for comorbidity incidence. These findings are
plotted in supplementary Figure S2. For example, epilepsy patients under the age of 64 years showed the highest incidence of migraine and over the age of 64 years had the highest incidence of dementia [33] — both of which are inconsistent with the trend observed in literature co-occurrence. Interestingly, neither of these trends correlated with the number of disease-specific literature, which reflects publication bias.

<table>
<thead>
<tr>
<th>Subgraph name</th>
<th>Biological entities</th>
<th>Relationships</th>
<th>Components</th>
<th>Citations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adaptive immune system subgraph</td>
<td>12</td>
<td>12</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Adenosine signaling subgraph</td>
<td>76</td>
<td>154</td>
<td>3</td>
<td>15</td>
</tr>
<tr>
<td>Apoptosis signaling subgraph</td>
<td>228</td>
<td>503</td>
<td>5</td>
<td>115</td>
</tr>
<tr>
<td>Brain-derived neurotrophic factor signaling subgraph</td>
<td>75</td>
<td>142</td>
<td>1</td>
<td>29</td>
</tr>
<tr>
<td>Calcium dependent subgraph</td>
<td>302</td>
<td>793</td>
<td>8</td>
<td>73</td>
</tr>
<tr>
<td>Chromatin organization subgraph</td>
<td>8</td>
<td>10</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Energy metabolic subgraph</td>
<td>91</td>
<td>177</td>
<td>4</td>
<td>24</td>
</tr>
<tr>
<td>Estradiol metabolism</td>
<td>7</td>
<td>8</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>G-protein-mediated signaling</td>
<td>78</td>
<td>140</td>
<td>5</td>
<td>25</td>
</tr>
<tr>
<td>Gaba subgraph</td>
<td>262</td>
<td>632</td>
<td>2</td>
<td>56</td>
</tr>
<tr>
<td>Glutamatergic subgraph</td>
<td>121</td>
<td>246</td>
<td>5</td>
<td>32</td>
</tr>
<tr>
<td>Hormone signaling subgraph</td>
<td>126</td>
<td>256</td>
<td>9</td>
<td>16</td>
</tr>
<tr>
<td>Inflammatory response subgraph</td>
<td>42</td>
<td>49</td>
<td>4</td>
<td>21</td>
</tr>
<tr>
<td>Innate immune system subgraph</td>
<td>41</td>
<td>63</td>
<td>8</td>
<td>18</td>
</tr>
<tr>
<td>Interleukin signaling subgraph</td>
<td>46</td>
<td>111</td>
<td>3</td>
<td>17</td>
</tr>
<tr>
<td>Long term synaptic depression</td>
<td>64</td>
<td>129</td>
<td>2</td>
<td>21</td>
</tr>
<tr>
<td>Long term synaptic potentiation</td>
<td>125</td>
<td>252</td>
<td>2</td>
<td>46</td>
</tr>
<tr>
<td>Mapt-erk subgraph</td>
<td>313</td>
<td>706</td>
<td>5</td>
<td>68</td>
</tr>
<tr>
<td>Metabolism</td>
<td>340</td>
<td>600</td>
<td>14</td>
<td>126</td>
</tr>
<tr>
<td>Mira subgraph</td>
<td>5</td>
<td>4</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Mossy fiber subgraph</td>
<td>38</td>
<td>66</td>
<td>2</td>
<td>14</td>
</tr>
<tr>
<td>Mtor signaling subgraph</td>
<td>166</td>
<td>336</td>
<td>3</td>
<td>42</td>
</tr>
<tr>
<td>Neurtransmitter release subgraph</td>
<td>552</td>
<td>1667</td>
<td>5</td>
<td>131</td>
</tr>
<tr>
<td>Notch signaling subgraph</td>
<td>105</td>
<td>205</td>
<td>3</td>
<td>20</td>
</tr>
<tr>
<td>Protein kinase signaling subgraph</td>
<td>377</td>
<td>850</td>
<td>6</td>
<td>87</td>
</tr>
<tr>
<td>Protein metabolism</td>
<td>129</td>
<td>185</td>
<td>8</td>
<td>44</td>
</tr>
<tr>
<td>Reelin signaling subgraph</td>
<td>117</td>
<td>253</td>
<td>2</td>
<td>21</td>
</tr>
<tr>
<td>Regulation of actin cytoskeleton subgraph</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Serotonergic subgraph</td>
<td>148</td>
<td>478</td>
<td>2</td>
<td>18</td>
</tr>
<tr>
<td>Thyroid hormone signaling subgraph</td>
<td>106</td>
<td>228</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>Transport related subgraph</td>
<td>49</td>
<td>60</td>
<td>7</td>
<td>25</td>
</tr>
<tr>
<td>Wnt signaling subgraph</td>
<td>27</td>
<td>38</td>
<td>3</td>
<td>15</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>3478</strong></td>
<td><strong>12481</strong></td>
<td><strong>16</strong></td>
<td><strong>641</strong></td>
</tr>
</tbody>
</table>

**Table 2:** Statistics over the Epilepsy Knowledge Assembly, grouped by subgraph generated by PyBEL v0.10.0. The biological entities entry corresponds to the number of genes, chemicals, proteins, biological process, etc. in a subgraph. The number of relationships (i.e., edges) corresponds to the number of publications in which a relationship was described. The number of components corresponds to the number of "connected" graphs exist within a subgraph. Finally, the number of citations corresponds to the total number of articles that were curated in the subgraph. A more detailed summary is included in the supplementary information.
While gene-centric methods alone were insufficient for unraveling the shared pathophysiology in epilepsy and studied comorbidities, systematic approaches for identifying and evaluating shared mechanisms may explain the aggregate effects of their interactions that cannot be captured by simple measures like literature co-occurrence.

**Mechanism Enrichment**

Alzheimer's disease was chosen as a putative comorbidity of epilepsy for further investigation not only because it had the highest literature-based gene overlap with epilepsy (Table 1), but additionally because of the prior existence of the Alzheimer's Disease Knowledge Assembly [34] and its inclusion in NeuroMMSig.

Of the most frequently mentioned drugs in epilepsy literature (Figure S1), we identified carbamazepine as having both notable target representation in the Alzheimer's Disease Knowledge Assembly along with multiple references indicating its positive effects on memory and treatment of elderly patients with seizures [35] as well as positive effects in the treatment of AD patients [35,36].

In the following sections, we first use NeuroMMSig to investigate the mechanisms enriched by the targets of carbamazepine in the context of epilepsy. After, we make a comparative enrichment to identify possible overlapping mechanisms with AD in order to explore the therapeutic effects of the drug in both disease contexts.

**Epilepsy Mechanism Enrichment**

While carbamazepine has been observed to act through inhibition of sodium and calcium voltage-gated channels as well as activation of the GABA receptors [37, 38], its mechanism of action is still not fully understood [39]. In order to better understand the impact of the drug, the gene set of all of its known targets (Text S1), was queried on the NeuroMMSig mechanism enrichment server against the Epilepsy Knowledge Assembly.

Because the NeuroMMSig mechanism enrichment algorithm only provides relative scores, the top 10th percentile of results were used. Of the twelve networks with at least one mapped gene, those with an enrichment score in the top 10th percentile (above 0.696) were the adenosine signaling subgraph and the GABA subgraph (Table S1). While the increase of the purine nucleoside, adenosine, has been associated with the incidence of seizures, its mechanistic connection is still unknown [40]. However, recent research has identified several promising targets that regulate and balance adenosine levels such as adenosine kinase (ADK) and its receptors (ADORA family) [40, 41]. Similarly to adenosine, the inhibitory neurotransmitter response induced by GABA is responsible for balancing many excitatory signals occurring in the brain. Studies that investigated reduction and abnormalities in GABA-inhibitory processes lead to the development of GABA agonists (e.g., vigabatrin and tiagabine) that act as anticonvulsants in epilepsy patients [42].

After assessing the plausibility of these subgraphs' involvement in the aetiology of epilepsy, the union of the networks was used to further investigate the relation between the downstream effects of carbamazepine and its therapeutic effect on epilepsy patients. Due to the its size, it was necessary to filter and query the network by finding the shortest paths between the drug's targets and different biological processes of interest in the context of epilepsy then combining them to form a new graph to support quickly identifying candidate
pathways. Finally, the common upstream controllers between each pathway were included to provide further context to their overlap. Combining, reasoning over, and manually interpreting the generated paths lead to the simplified network depicted in Figure 2 representing the downstream effects of carbamazepine.

The figure demonstrates the synergistic effects of the activation of the GABA receptor family and the inhibition of the sodium voltage-gated channels to potentiate GABA-mediated inhibition, and therefore, decrease the risk of developing seizures. Furthermore, carbamazepine causes an increase in the production of adenosine, whose receptors, ADORA1 and ADORA3, in a positive feedback loop with the production of adenosine which ultimately leads to a reduction in glutamatergic excitatory signals. Synaptic plasticity in response to the downstream effects of these signals may contribute to drug resistance, and ultimately, seizure recurrence.

Figure 2: A schematic representation of the knowledge surrounding carbamazepine retrieved by querying its targets with NeuroMMSig. The relevant portions of the most significantly enriched graphs in the context of epilepsy and AD, the adenosine signalling and the GABA subgraphs, were merged and displayed in order to highlight a potential explanatory mechanisms for the therapeutic effects of carbamazepine (Text S2). It is rendered with a hierarchical layout to mirror the flow from molecular entities to proteins, biological processes, and pathologies.
**Comparative Mechanism Enrichment**

NeuroMMSig was queried with the same gene set in the context of AD in order to perform a comparative investigation of shared mechanistic perturbations with epilepsy. Because the mechanism of action of carbamazepine is poorly represented in the literature (Table S2) and its known targets are less implicated in AD, it was unsurprising that fewer subgraphs were enriched in the context of AD.

The most significant, the GABA subgraph, which describes the upstream controllers of the GABA receptor and its downstream effectors, highly overlapped with the GABA subgraph in the context of the epilepsy - it contains key relations that may explain the efficacy of carbamazepine in both conditions. Studies in AD models have shown a negative correlation between the abundance of amyloid beta 42 (Aβ42) and the expression of the transcription factor for the GABA receptor family EGR1 [43,44]. This correlation could be caused by an unknown controller in the aetiology of AD; in which state, a patient would have decreased expression of EGR1, and therefore fewer GABA receptors and less ability to inhibit the excitatory signals that lead to seizures. While the link between epilepsy and GABAergic neurotransmission has already been exploited by antiepileptic drugs, its role in AD is not yet clearly understood (Figure 2). However, several recent publications have rationalized targeting GABAergic neurotransmission for treatment of AD [45, 46].

Tangentially, EGR1 upregulates the expression of PSEN2, a member of the catalytic subunit of the γ-secretase complex that regulates APP cleavage [47]. Mutations in PSEN2 that have been both linked to amyloid beta accumulation [48] and seizures in AD patients [49] provide further evidence for the existence for a shared mechanism through which carbamazepine acts in the contexts of AD and epilepsy (Figure 2).

Noradrenaline is a known anticonvulsant [50] that is often lacking in patients in the early stages of AD due to an observed loss of noradrenergic neurons [51]. Because carbamazepine has been observed to activate noradrenergic neurons [52], its anticonvulsant activity may be due to it indirectly increasing noradrenaline levels. Finally, noradrenaline potentiates the previously mentioned adenosine pathways [53].

While mechanism enrichment provides several insights, a cursory search of PubMed of "Carbamazepine"[nm] AND "Alzheimer disease"[MeSH Terms] suggested publications [54,55] that implicate autophagy in the therapeutic action of Carbamazepine. Further investigation showed Carbamazepine does not have a significant effect on expression of proteins in the mTOR-pathway [54] and that it is likely increasing autophagic flux through an mTOR-independent pathway [55]. Mechanism enrichment analysis was unable to prioritize autophagy pathways due to some of the shortcomings of knowledge-based methods. For example, the AD NeuroMMSig subgraph corresponding to autophagy pathways did not contain any of the targets of Carbamazepine listed by PharmGKB. This could be due to the choice of boundaries in subgraph definition, or also due to the lack of annotation of autophagy-related targets in PharmGKB. Autophagy has been implicated in epilepsy [56], but the literature has not yet succinctly described the connection from the therapeutic to its target, pathway, and finally the pathology. Finally, because knowledge assemblies are inherently incomplete, this shows the complementary nature of the two approaches.

While the exact mechanism of action of carbamazepine remains elusive, this proposed mechanism enrichment approach was able to identify multiple pathways through which it could be acting in both the AD,
epilepsy, and shared context. Looking forward, this approach can be applied across a wide variety of chemical matter in the neurodegenerative disease space, as well as in other domains for which appropriately annotated knowledge assemblies exist, in order to support identification of drugs' mechanisms of actions, drug repositioning opportunities, and the development of new lead compounds.

**Conclusion**

Our findings indicate that literature-based methods as a proxy for genetic pleiotropy generally do not correlate with the results from epidemiological studies of epilepsy and its comorbidities. Furthermore, strictly gene-centric methods lack the ability to elucidate mechanistic insight that a knowledge assembly can support.

After formalizing a representative sample of the knowledge surrounding epilepsy, its risk factors, its comorbidities, and antiepileptic drugs, we annotated mechanistic subgraphs to include in the NeuroMMSig mechanism enrichment server. Finally, an enrichment approach focusing on the targets of carbamazepine proposed the several downstream effectors (including the GABA receptor, GABAergic pathways, etc.) that could explain its therapeutic effects in both the contexts of epilepsy and AD.

Future work will include applying this procedure to a wider variety of drugs and chemical matter across different diseases. Finally, we have made the Epilepsy Knowledge Assembly publicly available through NeuroMMSig ([http://neurommsig.scai.fraunhofer.de](http://neurommsig.scai.fraunhofer.de)) to facilitate further systems biology and chemoinformatics investigations of epilepsy.

**Authors’ Contributions**

C.T.H. and D.D.F. conceived and designed the study, implemented and interpreted the mechanism enrichment approach in NeuroMMSig, and wrote this manuscript. N.B. and A.G curated and annotated the Epilepsy Knowledge Assembly. M.H.A. reviewed the content.

**Acknowledgements**

The authors would like to thank our colleagues at Fraunhofer SCAI for their critical review of this manuscript.

**Funding**

This work was supported by the European Union/European Federation of Pharmaceutical Industries and Associations (EFPIA) Innovative Medicines Initiative Joint Undertaking under AETIONOMY [Joint Technology Initiatives grant number 115568], resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies in kind contribution.

**Competing Interests**

None declared.
References


of excitatory synaptic transmission in area CA1 of the rat hippocampus. European Journal of Neuroscience 20: 1555-1565.

