

Whole exome sequencing studies in epilepsy: a deep analysis of the published literature

AUTHOR(S)

Arif Shukralla, Robert Carton, Katherine Benson, Hany El-Naggar, Austin Lacey, Gianpiero Cavalleri, Norman Delanty

CITATION

Shukralla, Arif; Carton, Robert; Benson, Katherine; El-Naggar, Hany; Lacey, Austin; Cavalleri, Gianpiero; et al. (2022): Whole exome sequencing studies in epilepsy: a deep analysis of the published literature. Royal College of Surgeons in Ireland. Journal contribution. <https://hdl.handle.net/10779/rcsi.19267880.v1>

HANDLE

[10779/rcsi.19267880.v1](https://hdl.handle.net/10779/rcsi.19267880.v1)

LICENCE

CC BY-NC-SA 4.0

This work is made available under the above open licence by RCSI and has been printed from <https://repository.rcsi.com>. For more information please contact repository@rcsi.com

URL

https://repository.rcsi.com/articles/journal_contribution/Whole_exome_sequencing_studies_in_epilepsy_a_deep_analysis_of_the_published_literature/19267880/1

Whole Exome Sequencing Studies in Epilepsy: A Deep Analysis of the Published Literature.

Arif Shukralla^{1*}, Robert Carton^{2,3,4}, Katherine A. Benson^{2,3,4}, Hany El Nagggar^{1,2,3}, Austin Lacey^{2,3}, Gianpiero Cavalleri^{2,3,4}, Norman Delanty^{1,2,3}

¹ The National Epilepsy Programme, Beaumont Hospital, Dublin, Ireland

² FutureNeuro, The SFI Research Centre for Chronic and Rare Neurological Disease, Ireland

³ The Royal College of Surgeons in Ireland, Dublin, Ireland

⁴ School of Pharmacy and Biomolecular Science, RCSI, Dublin, Ireland

* Corresponding Author

Abstract

Purpose: To evaluate the quality of Whole Exome Sequencing (WES) reporting in the epilepsy literature.

Method: We aimed to assess the quality of reporting of WES in epilepsy. We compared studies based on journal type and if outcome reporting biases exist. We used a self-constructed benchmark to quantitatively analyse studies.

Results: We included 451 publications. Reporting was heterogeneous with poor reporting of 1) ACMG guideline application 13% and 2) Human Phenotype Ontology (HPO) numbers in 3% of studies, 3) VUS in 19%. Predictors of reporting included journal type and journal impact factor. Date of publication and publication type were not predictors of poor reporting. Pairwise comparisons of genetics versus neurology journals using relative risks yielded significant differences in reporting of ACMG guideline application (RR 1.88 CI 1.04-3.38); HPO numbers (RR 8.62 CI 1.08-63.37) and deposition of findings to ClinVar (RR 2.50 CI 1.03-6.1).

Conclusion: Reporting of WES literature is heterogeneous in quality, and poor reporting hinders collaboration and accession of data into large databases like OMIM and OrphaNet. This study highlights reporting bias in this area and, formal structural guidelines like the CONSORT guidelines used in the reporting of clinical trials are needed to address the issue.

Keywords

Whole exome sequencing, CONSORT, reporting standards, epilepsy

Introduction

Epilepsy is a condition that predisposes patients to recurrent seizures. The diagnosis of epilepsy requires the presence of two or more seizures or one seizure with the predisposition to further seizures evidenced by medical tests or the diagnosis of an epilepsy syndrome (Fisher et al. 2014). In the latter case, either a single gene mutation or several gene mutations are the cause in approximately 70% of patients (Myers & Mefford, 2015). Knowledge of the genetics of epilepsy has undergone significant advances, and genetics now permeates most of the epilepsy research. Specific epilepsy syndromes, such as Dravet syndrome, a known monogenic cause with an associated set of causal genes, and single-gene testing may still be appropriate in these situations. Although targeted gene panel sequencing is still the first-line test, whole-exome sequencing (WES) is cost-effective and more appropriate in many situations where the cause of epilepsy is not clear, such as in developmental epileptic encephalopathies and non-lesional focal epilepsies.

WES is a method for rapidly delineating the genome's protein-coding sequence, namely exons, to identify pathogenic variants (Mefford 2012). Sequencing the proband and both parents, using trio WES, can further determine if a variant was inherited from either parent or arose *de novo*. WES is cost-effective in situations such as developmental epileptic encephalopathies, providing a diagnostic yield of 33% and can change clinical management in 39% of patients (Demos et al., 2019). WES is replacing gene panel testing due to the falling costs of sequencing. Reading, aligning, and interpretation of sequencing output is fraught with bias. Biases may

arise in the methodological process of variant calling, variant filtering, or results interpretation. The literature provides some recommendations to address these biases (Richards et al. 2015).

Epilepsy-related research facilitated by WES has significantly increased since 2012 and, most of the published literature is in the form of case reports, case series or cohort studies. Although American College of Medical Genetics (ACMG) guidelines exist for methodological issues and patient consent, to date, there are no guidelines on how authors should report diagnostic sequencing in the literature (Richards et al. 2015). As founded on the principles of open disclosure in clinical trials and case reports, the CONSORT group has published a set of guidelines for reporting these studies (Schultz, Altman & Moher, 2010). Their mission is to ensure the quality of trials reports, and the CONSORT documents are frequently used as a benchmark for audit and research purposes.

In this analysis, we seek to explore the reporting of studies of epilepsy patients undergoing WES. We used a self-constructed benchmark to analyse studies quantitatively. Our primary aim was to assess the quality of reporting of WES. Secondary aims were to compare the quality of studies based on journal type, to explore how reporting has changed over time, and determine the presence of any reporting biases-

Methods

Eligibility and identification of studies

We used MEDLINE to search for citations of studies using search terms "epilepsy" and "whole exome sequencing" between 2011 and the end of 2019. Inclusion criteria were studies of patients with epilepsy, using WES to diagnose or confirm the epilepsy syndrome. We excluded studies that did not use WES, such as whole-genome sequencing, copy number variant studies, studies not in English and studies that were secondary reports of earlier studies. We retained studies with epilepsy, intellectual disability or autism if the patient/patients' phenotype included seizures. We excluded studies where seizures were not the dominant phenotype. Studies were included by reading the abstract and, where necessary, reading the full text. Two authors, A.S. and R.C., collaborated in selecting studies for inclusion and any disparity discussed at *ad-hoc* meetings.

Data collected

There are no published guidelines on the reporting of WES studies. But, there are several guidelines for recommendations on reported variant calling, methodological soundness of variant calling, and disclosing WES results to family members and clinicians-(Matthijs et al. 2016; Green et al. 2013; Hehir-Kwa et al. 2015). We compiled a checklist of items pertaining to sections of a typical paper reporting WES results. Table 1 shows selected items with comments on why they were included. A review team (A.S. R.C. N.D. G.C. and H.EN.) discussed all items for suitability and relevance for this study. Of the 38 items in Table 1, items 10 and 29 apply to cohort studies only. We also collected journal name, publication date and

year, journal impact factor (used as a surrogate for audience size), and type of study (case vs cohort).

Data extraction

We devised a data extraction form; two authors (A.S. and R.C.) independently extracted data from a random sample of 15 studies. We compared results to determine inter-rater reliability. Any disagreements were clarified by mutual discussion, and if data extraction was deemed reliable, both authors completed the data extraction process.

Data Analyses

We calculated the inter-rater agreement of extracted data using Cohen's Kappa to measure agreement (Cohen 1968). For each item met in the checklist, we awarded one point. Hence, the minimum score possible for case studies or case series is zero, and the maximum is thirty-six. For cohort studies, we added two additional items, so the maximum possible score is thirty-eight. For between-group comparisons, we used an independent-pair *t-test*. We used a one-way ANOVA and Tukey's Honest Significant Difference test to conduct *post hoc* analyses to control for multiple comparisons (Tukey 1949). The proportion of studies reporting each item of interest calculated, and relative risks, with 95% confidence intervals, were used to summarise comparative effects. We used SPSS version 28 to conduct statistical analyses using a two-sided significance level of 5% and relative risks calculated using RevMan version 5.4.

Results

Study selection and inclusion

One author, A.S., selected studies. We found 920 citations of WES and epilepsy published between 2011 and 2019 (Figure 1). Of these, we excluded studies not in English (14 studies), studies not on epilepsy (216 studies), and we excluded 101 review articles and 103 studies not using whole-exome sequencing. A total of 451 studies were thus selected for analysis. Included studies are listed in appendix A.

Inter-rater agreement for data extraction

We calculated Cohen's kappa to determine the inter-rater agreement for data extraction (Cohen 1968). Two authors, A.S. and R.C., extracted data from 23 studies. A kappa score of 0.61 was obtained with a percentage agreement of 80.8%, indicating substantial agreement between authors, and therefore data extraction was deemed reliable (Glen S. 2014).

Characteristics of studies

We categorised studies by syndrome, journal type, study methodology, publication date, and journal impact factor.

Developmental epileptic encephalopathy comprised 33% of all studies, epilepsy with intellectual disability comprised 24% of studies, with the remainder of 43% of studies consisting of other syndromes. These syndromes included focal-onset epilepsies, focal cortical dysplasia, progressive myoclonic epilepsies and other syndromes. Fifty per cent of studies were published in genetics journals, and thirty-seven per cent of studies were published in neurology journals; thirteen per cent of studies were published in other journal types (Table 2). The top ten publishing journals are shown in Table 3, indicating that the American Journal of Medical Genetics had the most selected studies (7.3%) (Table 3).

Case reports or case series comprised a large proportion of the studies (80%), whereas cohort studies consisted of 20%. We chose a date of the end of 2016 to compare reporting trends before and after this date. This was arbitrarily set to ensure roughly equal datasets. Two hundred and eighteen studies were published before 2017, and two hundred and thirty-three studies were published after 31st December 2016 (Table 2). One hundred and eighty-eight studies were published in low impact factor journals (42%) and seventy-six studies in high impact factor journals (13%). We categorised journals with an impact factor of three or less as low impact factor journals and those with eight or above as high impact factor journals.

Quality of reporting

We calculated a quality score totalling a maximum of 36 for case studies and 38 for cohort studies. The mean score across all studies was 21.2 (95% confidence

intervals (CI) of means 20.7 to 21.6), a median score of 22 with a range of 4 to 31 (Table 4). No study scored a maximum of 38 points. We carried out Tukey's Honestly Significant Difference (HSD) post hoc analyses showing significant differences in means for genetics and neurology journals; genetics and other journals but not between neurology and other journals (Appendix B)

We categorised items into those reported well, those reported with average frequency and those reported poorly (Table 5). The reporting of WES data in epilepsy is heterogeneous. Items that were reported poorly include Human Phenotype Ontology (3%), deposition of results to ClinVar (6%) (Landrum et al. 2018), whether ACMG guidelines were applied (13%) and quality control measures used in the Methods section (14%). Other items that were reported poorly included the reporting of variants of undetermined significance (19%).

Items reported with average frequency included Online Mendelian Inheritance in Man (OMIM) numbers (35%), previous genetics tests carried out (35%), reporting of annotation databases (38%), and reference to evolutionarily conserved databases (31%). Ethnicity was reported in 61% of studies, and consanguinity was reported in 58%. At least one pedigree chart was presented in 53% of studies, DNA source was reported in 47% of studies, reference genome to assemble exome data was reported in 45% of studies and types of *in-silico* prediction tools were reported in 59%. How variant impacts the function of the polypeptide was reported in 64% of articles. A diagnostic yield was reported in 43% of studies, and the full transcript of information was reported in 47% of studies. The mutated variant or variants were mentioned in 63% of studies.

The items that were reported well included: gene mentioned in abstract (97%), syndrome or clinical features mentioned in the abstract (96%), clinical features mentioned in the text (95%), mutation type (95%), variant inheritance/segregation (91%), discussion of the previous literature (98%), ancillary investigations (88%), and phenotype comparison to associated diseases (86%).

Genetics versus Neurology Journals

Genetics journals scored 22.4 points, and neurology journals scored 19.9 points with a mean difference of 2.5 (95% CI of means 1 to 3.5), see Table 4. This was statistically significant, with a *p*-value of <0.01, indicating that genetics journals report superior WES data based on the criteria of this study.

When comparing individual items, we examined relative risks (R.R.) shown in figures 2 to 7 with all data shown in appendix C and D. Relative risks of items whose confidence intervals did not include unity are statistically significant and therefore worthy of note. For brevity, confidence intervals are shown in the figures only and not in the text here. Our findings showed a consistent trend of relative risks favouring genetics journals (figure 2). Genetics journals applied the ACMG guidelines almost twice as frequently as neurology journals (RR 1.88). Items reported better in genetics journals were Human Phenotype Ontology (RR 8.62); and OMIM data (RR 1.88). Genetics journals were better in reporting items on bioinformatics methodology; these items include exome target kit (RR 1.26),

type of sequencer used (RR 1.14), quality control data (RR 1.95), reference genome used assemble exome data (RR 1.57), haplotype caller used (RR 1.24), variant filtering (RR 1.26), minor allele frequency databases (RR 1.19), evolutionarily conserved databases consulted (RR 1.54) and how variant impacts translated polypeptide (RR 1.32).

When discussing results, genetics journals reported Sanger sequence validation (RR 1.14) and reported the full transcript information (RR 1.41) compared to neurology journals. Genetics journals were more likely to report sequence chronogram data (RR 1.67) and were more likely to deposit their findings to ClinVar (RR 2.5).

Genetics versus Other journals

Genetics journals scored 2.3 points higher than other journals (95% CI of means 1 to 3.5; $p < 0.01$). Genetic journals reported OMIM (RR 2.05) and consanguinity data (RR 1.36) better than other journals (figure 3). Also, prior genetics tests (RR 1.73) were better reported in genetics journals. In the methods section, genetics journals reported exome target kit (RR 1.3), sequencer type (RR 1.25), reference genome used to assemble exome data (RR 1.46), variant filtering approach (RR 1.47), minor allele frequency (RR 1.3) and *in-silico* prediction tools (RR 1.35) better than other journals. Diagnostic yield was reported more in other journals compared to genetics journals (RR 0.53). Full transcript information was documented better in genetics journals (RR 1.52). Phenotype discussion was more complete in genetics journals (RR 1.19).

Neurology versus Other journals

Neurology journals scored 19.9 points, and other journals scored 20.1 points. There was no significant difference between total scores of neurology and other journals with a mean difference of 0.2 (95% CI of means -1.7 to 1.3) (Table 4). Only six items were statistically significant comparing neurology with other journals (figure 4). Neurology journals were better than other journals in discussing previous literature (RR 1.60), and the phenotype of the patients (RR 1.21). However, neurology journals performed poorly in reporting DNA source (RR 0.66), quality control measures (RR 0.46), mutation type (RR 0.94) and sequence chronogram (RR 0.94).

Studies published before and after January 2016

There was no significant difference between studies published before or after the end of 2016. Pre- 2017 studies scored 20.9 points, and post-end 2016 studies scored 21.4 points with a mean difference of 0.5 (95% CI of means -1.4 to 0.4). Items that were reported well before the end of 2016 include gene name in title (RR 0.92) (figure 5), exome target kit (RR 0.85), how variant impacts polypeptide (RR 0.8). Items that were reported well after 2017 were molecular science background (RR 1.19); *in-silico* prediction tools (RR 1.19); pathogenic databases of previous reports (RR 1.29) and ACMG guidelines (RR 12.63). Other items reported well after 2017 were: full transcript (RR 1.45); journals depositing data to ClinVar (RR 2.90) and DNA source (RR 1.24).

Case vs Cohort Studies

We found no difference in scores between the case and cohort studies. Case studies scored 21.17 points, and cohort studies scored 21.08 points with a mean difference of 0.5 (95% CI -1.4 to 0.4). However, some items are better reported in cases studies. These items included gene named in the title (RR 1.39) (figure 6), abstract mentions variant (RR 2.22), and abstract mentions gene (RR 1.11). Clinical features (RR 1.13) and consanguinity (RR 1.74) were reported better in case of studies than cohort studies, ancillary investigations (1.15) and pedigree charts (RR 1.69) were reported better in case studies. Sequence chronogram (RR 1.98) and phenotype compared to reported gene (RR 1.16) were reported better in case studies. Cohort studies reported OMIM data (RR 0.71), prior genetic tests (0.72), haplotype caller (RR 0.79), variant filtering (RR 0.8) and annotation databases used (RR 0.77) better than case studies. ACMG guidelines, if applied, were better reported better in cohort studies (RR 0.24), and variants of undetermined significance (RR 0.57) were included more in cohort studies.

Impact factor

High impact factor journals scored 22 points, and low impact factor scored 20.6 points with a mean difference of 1.4 (95% CI of means -2.7 to -0.05). Items that were better reported in high impact factor journals included OMIM (RR 1.47) (figure 7), reference genome to assemble exome data (RR 1.39), haplotype caller (RR 1.30), variant filtering approaches (RR 1.5) and annotation databases (RR

1.40). High impact factor journals reported how variant impacts polypeptide function better than low impact journals (RR 1.72). High impact journals reported the full transcript of variants better than low impact factor journals (RR 1.51) (figure 7).

Items that were reported better in low impact factors journals are clinical features (RR 0.89), diagnostic yield (RR 0.50) and the phenotype description compared to the reported disease (RR 0.75). Variants mentioned in the abstract was better reported in low impact journals (RR 0.79).

High impact genetics journals versus low impact genetics journals

High impact genetics journals reported the variant in the abstract more than low impact journals (RR 1.44). Also, high impact genetics journals reported clinical features better than low impact journals (RR 1.18) (Appendix D). Inclusion criteria were reported better in high impact journals (RR 2.10); however, low impact genetics journals reported how variant affects polypeptide structure better than high impact genetics journals (RR 0.77).

High impact genetics journals versus high impact neurology journals

We found that ethnicity and OMIM numbers were reported better in high impact genetics journals compared to high impact neurology journals (RR 2.39) and (RR 2.39), respectively (Appendix D). We also found that neurology journals were

better at reporting variant segregation or inheritance than genetics journals (RR 0.85).

Discussion

Reporting standards of whole-exome sequencing is a novel area of research in epilepsy. Here, we carried out a quantitative review of the quality of WES studies in the epilepsy literature. Our results showed that the reporting of WES data is heterogeneous, with some items reported well and others poorly. These reporting trends mirror our findings in the area of adverse events of antiepileptic drugs, where the reporting emphasis resides in the introduction and results sections at the expense of methodology (Shukralla et al.–2011). Similarly, we found that reporting quality varies in the methods, results and discussion sections of publications.

Reliability is whether a given test or method is reproducible. WES output reliability relates to how sequencing is performed, the quality of WES reads and their related bioinformatic scores. Validity is a measure of how a method measures what it is intended to measure. Validity is determined by the laboratory's accreditation and coverage consistency (Gotway et al., 2020). We know from the statistical methodology literature that inadequate reporting of clinical trials and other literature has implications on both the reliability and validity of systematic reviews (Liberati et al. 2009); and we would argue that poor reporting of WES research data will ultimately have ramifications on how data is synthesised in large databanks such as ClinVar, OMIM, ExAC and the NLHBI exome

project, as they rely on provider input. Indeed, earlier work highlighted inaccuracies in variant interpretation that led to forming a central database such as ClinVar (Harrison et al. 2017). Essential items such as HPO terms, evolutionary conserved databases and full transcripts need further elaboration.

The quality of pathogenicity rating for variants is linked to adherence to ACMG guidelines, and we found that only 13% of studies reported using the guidelines in their analyses. We found that genetics journals reported ACMG guidelines and their implementation better than neurology journals, reflecting its editorial standards. We propose that variants of undetermined significance are published in the literature to inform clinicians that a given variant is unlikely to be pathogenic and prevent the error of attributing a phenotype to a variant that may not be relevant and avoid errors in misclassification. Nevertheless, variants may still prove to be pathogenic. In any case the ACMG guidelines need to be adhered to and one isn't sure if this is pervasive in the literature as it stands. The ClinGen criteria are curated by the ACMG guide variant assessment. In this review, we did not examine ClinGen criteria per se, which is a limitation of this study.

Human Phenotype Ontology (HPO) is a novel way of categorising and comparing phenotypes in genetics studies; HPO allows for deep phenotyping and is used in WES diagnosis (Robinson et al. 2008; Köhler et al. 2019). The nervous system accounts for the second most common terms in the HPO database, with epilepsy and intellectual disability contributing to most neurological conditions (Köhler et al., 2020). We found that only a mere 3% of articles mention HPO numbers. As of writing, the American Journal of Medical Genetics requires authors to include HPO

terms; however, other non-genetic journals do not instruct authors to include HPO terms. This may be due to a lack of awareness of the HPO system and how it may help standardisation and phenotyping. The OrphaNet database of rare diseases uses HPO terms to collect individual disease information from the literature and provides unique disease identifier codes and epidemiological data. If HPO terms are not reported, rare diseases are difficult to phenotype and not able to be inputted into this valuable resource.

Our analysis found that OMIM number was reported in 35% of studies. OMIM was reported more consistently in genetics journals than in neurology and other journals; interestingly, OMIM was better reported in low impact factor journals and better reported in case studies. Databases such as ClinVar accept submissions from external sources based on phenotypic correlation, a poor reporting of OMIM in the literature will delay the ability of databanks in capturing new data (Landrum et al., 2018). ClinVar recommends the submission to contain either OMIM, OrphaNet, HPO or other numbers (NCBI.NLM.NIH, 2020).

Evolutionary conserved databases are frequently consulted to help confer pathogenicity to a given variant. If a variant is located within an evolutionary conserved region designated by lineage conservation databases, it is further evidence that the variant may be pathogenic. Less than a third of the studies we analysed referenced these databases. Furthermore, they did not or provide a chronogram of the variants across species. We found that genetics journals reported evolutionary conserved databases better than neurology journals, indicating a lack of awareness and potential skill differences in authorship.

Quality control measures such as Phred score or Q-score and other data was reported in just 14% of studies. These scores provide a quality metric of genetic data produced from an NGS; for example, a low Phred reduces the probability that a particular base has been called correctly, therefore our confidence that a variant is genuine. We found that quality scores were two times more likely to be reported in genetics journals than neurology journals. Poor reporting of quality scores is commensurate to incomplete reporting leading to bias in the findings' significance (Kong et al. 2018).

Ethnicity, consanguinity and pedigree charts are data elements helpful in interpreting autosomal recessive conditions; also, they communicate segregation of variants in a meaningful way. Ethnicity is essential for interpreting variants as this will vary depending on the ethnic background (Petrovski & Goldstein, 2016). Our study demonstrated that nearly half of the articles published still do not report them. We found no differences in our subsets regarding ethnicity, but consanguinity was reported better in genetics journals and case studies. Furthermore, our analyses showed that genetics journals were twice as likely to report ethnicity than neurology journals within the subgroup of high impact factor journals.

Most articles identified DNA sources derived from blood; other DNA sources include saliva or organ tissue. WES needs high-quality DNA free from contaminants. One study in cancer genomics found that irrespective of DNA source, WES results were little affected (Zhu et al., 2015). However, blood is

recommended with whole-genome sequencing as other DNA sources (e.g., saliva) are less efficient for sequencing genomic DNA or have a slightly elevated false-positive rate (Troost et al., 2019). Nevertheless, if DNA source is not reported, comparative assessments between studies would need to be interpreted with caution. Our analysis showed that the DNA source was reported in less than half of the studies. Reporting DNA sources does not consume significant space in a journal article and reflects that lack of guidance to authors in this regard.

Only 19% of studies we examined reported variants of undetermined significance (VUS). VUS reporting has minor precision medicine implications; nevertheless, poor reporting indicates reporting bias. We found that poor reporting of VUS in the medical literature is a cause of concern; this limits the accumulation of a comprehensive database of variants whose significance may be appreciated at a future period. In the era of ever-increasing publication of WES in epilepsy, and one expects the incidence of VUS to increase with increasing annotation data (Dunn et al., 2018). We found that cohort studies are two times more likely to report VUS than case studies. We postulate that cohort studies sequence a larger number of patients and are more likely to be published in genetics journals, accounting for the trend.

Reporting the full transcript of the variant is essential as it allows other researchers to verify the variant. Our study showed that a full transcript was reported in 47% of all studies. A full transcript makes a variant publicly available, and this was reported better in genetics journals than in neurology and other journals. Case studies and studies published after 2016 were more likely to

disclose the full transcript of the variant. Given that most WES literature is published as case studies, it is vital that the full transcript be available for reproducibility and disclosure.

Genetics journals reported WES data better than neurology journals, and this was statistically significant. This is likely due to these journals' readership, authorship and editorial requirements; clinical authors would prefer neurology journals for various reasons, and peer review may overlook the finer points of bioinformatic data reporting. Essential steps in bioinformatic methodology are reported better in genetics journals, including exome target kit, type of sequencer, quality control measures and change from the reference genome. Results such as how variants impact polypeptide, full transcript information, sequence chronogram and deposition of results to ClinVar were reported better in genetics journals.

Neurology journals performed poorly compared to other non-genetic journals in reporting overall. Other journals were better at reporting quality control measures such as a Phred score, and other journals reported DNA sources better than neurology journals. Neurology journals were relatively poorer at reporting the type of mutation and often did not include a sequence chronogram in the results. However, neurology journals were more likely to correlate phenotype to previously described disease and discuss the previous literature. Overall, the comparison between neurology and other journals highlights that neurology journals are significantly poorer at reporting WES methods, results and outcomes.

Case studies form the bulk of WES literature, forming roughly 80% of all studies. Although reporting standards exist for case studies (von Elm et al., 2007), these do not apply to WES methodology or results. Case reports, by their nature, reported some items better than cohort studies, including abstract data, clinical features, consanguinity and pedigree charts. However, this was at a cost to poor reporting of methodological data, variants of undetermined significance (VUS), and application of ACMG guidelines.

This study highlights the deficiencies in neurology journals which include epilepsy journals in the reporting of WES data. We did not find any specific instructions to authors concerning genetic data in epilepsy journals. Most journals mention the CONSORT statements in their instructions, yet, there are no guidelines on reporting WES. The CONSORT statements cover case reports and reporting standards for genome-wide associations studies (GWAS). However, GWAS studies are rarely performed, and updated guidance is needed in the era of WES studies in epilepsy.

Surprisingly, overall, we did not find any improvement in WES reporting over time. However, reporting of ACMG guidelines improved after 2016, and studies were more likely to report the full transcript and studies were more likely to deposit their findings to ClinVar. Reporting of some items did not improve after 2016 and these included a discussion of how the variant impacts polypeptide and exome target kit.

WES is a novel diagnostic methodology with the means of uncovering disease causative genes. However, the use of WES is influenced by correct methodology and variant interpretation often, with a multidisciplinary approach. Therefore this study highlights gaps in reporting and should prompt further discussion on how these gaps are addressed. We propose a joint task force of the ILAE and the CONSORT to issue guidance on how next-generation sequencing data is published in epilepsy and medical specialties literature.

Precision medicine therapies are now beginning to evolve for the treatment of the epilepsies, underscoring the increasing need and utility of precision diagnostics with WES (Baldassari et al. 2016). As WES becomes more cost-effective, it will enter more routine and appropriate clinical practice. Clinicians will need instruction and guidance in reading the ever-increasing literature in this field, and thus the quality of reports needs to be standardised across journals. Though many genetics journals already have standards for reporting variants and genomic data, these standards need to be cited in other specialty journals as genomic methods become mainstream. Journal editors need to review instructions to authors and how they publish genetic data to enhance transparency and allow for comparative assessments.

This study is limited in its scope by using quantitative data to assess the quality of reporting. Although this study explored the reporting of the literature, further work is needed to review if authors use specific standards. The strength of this study is the large sample size used in the analysis. In the absence of a suitable benchmark, we compiled our benchmark, which underwent a collaborative phase

before implementation in our study. We would welcome others to comment and add to this benchmark. Notwithstanding the deficiency, we highlighted multiple areas of poor reporting across WES studies in the epilepsies.

Recommendations

1. Based on the findings of this study, we make the following recommendations. Authors seeking to publish WES data should ensure that all relevant methodological detail is presented.
2. We recommend that authors be familiar with the ACMG guidelines and adhere to them. We recommend that journal editors be mindful of recent advances and instruct authors to submit articles with correct terminology.
3. A collaboration of the ILAE and the CONSORT groups should make new recommendations reporting WES in the literature. Allowing frequent updates as and when sequencing technology changes in the future.
4. A collaboration with the ILAE and HPO ontology should consider harmonising the ILAE classification with HPO terms.
5. Authors should be particularly careful in presenting phenotypic data, and we recommend using HPO terms.
6. Clinicians that outsource genetic analysis need to be aware of some of the methods used in variant calling and the significance of the output.

Disclosures

Authors GC and ND are in receipt of research funding from Congenica LTD, through the Science Foundation Ireland FutureNeuro Research Centre. This article was not funded.

We confirm that we have read the Journal's position on issues involved in ethical publication and affirm that this report is consistent with those guidelines.

References

Baldassari S., Licchetta L., Tinuper P., Bisulli F., Pippucci T. (2016). GATOR1 complex: the common genetic actor in focal epilepsies. *Journal of Medical Genetics*, 53, (8) 503-510. <http://dx.doi.org/10.1136/jmedgenet-2016-103883>

Cohen (1968). Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin*, 70, (4) 213-220
<https://doi.org/10.1037/h0026256>

Demos M., Guella I., DeGuzman C., McKenzie M.B., Buerki S.E., Evans D.E. ... Farrer M. (2019). Diagnostic yield and treatment impact of targeted exome sequencing in early-onset epilepsy. *Frontiers in Neurology*, 21, (10) 434
<https://doi.org/10.3389/fneur.2019.00434>

Dunn P., Albury C.L., Maksemous N., Benton M., Sutherland H.G., Smith R.A. ... Griffiths L. (2018). Next-generation sequencing methods for the diagnosis of

epilepsy syndromes. *Frontiers in Genetics*, 7, (9) 20

<https://doi.org/10.3389/fgene.2018.00020>

von Elm E., Altman D.G., Egger M., Pocock S.J., Gøtzsche P., Vandenbroucke J.P.

(2007). The Strengthening the Reporting of Observational Studies in

Epidemiology (STROBE) statement: guidelines for reporting observational

studies. *Annals of Internal Medicine*, 147, (8) 573-577

<https://doi.org/10.1016/j.jclinepi.2007.11.008>

Fisher, R.S., Acevedo, C., Arzimanoglou, A., Bogacz, A., Cross H., Elger C.E. ... Wiebe

S. (2014). ILAE Official Report: A practical clinical definition of epilepsy.

Epilepsia, 54, (4) 475-482. <https://doi.org/10.1111/epi.12550>

Glen S. (2014) What is Cohen's Kappa statistic? Retrieved from

<https://www.statisticshowto.com/cohens-kappa-statistic>

Gotway G., Crossley E., Kozlitina J., Xing C., Fan J. Hornbuckle C. ... Park J.Y.

(2020). Clinical exome studies have inconsistent coverage. *Clinical Chemistry*, 66,

(1) 199-206 <https://doi.org/10.1093/clinchem.2019.306795>

Green R.C., Berg J.S., Grody W.W., Kalia S.S., Korf B.R., Martin C.L. ... Biesecker L.G.

(2013). ACMG recommendation for reporting of incidental findings in clinical

exome and genome sequencing. *Genetics in Medicine*, 15, (7) 565-574

<https://doi.org/10.1038/gim.2013.73>

Harrison S.M., Dolinsky J.S., Knight-Johnson A.E., Pesaran T., Azzariti D.R., Bale S. ... Rehm H.L. (2017). Clinical laboratories collaborate to resolve differences in variant interpretation submitted to ClinVar. *Genetics in Medicine* 19, (10), 1096-1104 <https://doi.org/10.1038/gim.2017.14>

Hehir-Kwa J.Y., Claustres M., Hastings R.J., van Ravenswaaij-Arts C., Christenhusz G. Genuardi M. ... Robinson P.N. (2015). Towards a European consensus for reporting incidental findings during clinical NGS testing. *European Journal of Human Genetics Volume*, 23, 1601-1606 <https://doi.org/10.1038/ejhg.2015.111>

Köhler S., Carmody L., Vasilevsky N., Jacobsen J.O., Danis D., Gourdine JP., Gargano M. ... Robinson P.N. (2019). Expansion of the Human Phenotype Ontology knowledge base and resources. *Nucleic Acids Research*, 47 ,(D1) D1018-D1027 <https://doi.org/10.1093/nar/gky1105>

Köhler, S., Gargano, M., Matentzoglou, N., Carmody, L. C., Lewis-Smith, D., Vasilevsky, N. A., Danis, D., Balagura, G., Baynam, G., Brower, A. M., Callahan, T. J., Chute, C. G., Est, J. L., Galer, P. D., Ganesan, S., Griesse, M., Haimel, M., Pazmandi, J., Hanauer, M., ... Robinson, P. N. (2021). The Human Phenotype Ontology in 2021. *Nucleic Acids Research*, 49(D1), D1207–D1217. <https://doi.org/10.1093/nar/gkaa1043>

Kong S.W., Lee I.H., Liu X., Hirschhorn J.N., Mandl K.D. (2018). Measuring coverage and accuracy of whole-exome sequencing in clinical context. *Genetics in Medicine*, 20, (12) 1617-1626 <https://doi.org/10.1038/gim.2018.51>

Landrum M.J., Lee J.M., Benson M., Brown G.R., Chao C., Chitipiralla S. ... Maglott D.R. (2018). ClinVar: improving access to variant interpretation and supporting evidence. *Nucleic Acids Research*, 46, (D1) D1062-D1067.
<https://doi.org/10.1093/nar/gkx1153>

Liberati A., Altman D., Tetzlaff J., Mulrow C., Gøtzsche P.C., Ioannidis J.P. ... Moher D. (2009). The PRISMA statement for reporting systematic reviews and meta-analysis of studies that evaluate healthcare interventions: explanation and elaboration. *Journal of Clinical Epidemiology* 62 (10) e1-e34
<https://doi.org/10.1016/j.jclinepi.2009.06.006>

Matthijs G., Souche E., Alders M., Corveleyn A. Eck S. Fennstra I. ... Bauer P. (2016). Guidelines for diagnostic next-generation sequencing. *European Journal of Human Genetics*, 24, (1) 2-5 <https://doi.org/10.1038/ejhg.2015.226>

Mefford H.C. (2012). Diagnostic exam sequencing--are we there yet? *NEJM* 367 (20) 1951-1953 [10.1056/NEJMe1211659](https://doi.org/10.1056/NEJMe1211659)

Myers, C.T., Mefford, H.C. (2015) 'Advancing epilepsy genetics in the genomic era'. *Genome Medicine*, 7, (1) 91 <https://doi.org/10.1186/s13073-015-0214-7>

NCBI.NLM.NIH (2020) Instructions for ClinVar submission spreadsheets

Retrieved From

<https://www.ncbi.nlm.nih.gov/clinvar/docs/spreadsheet/#criteria>

Petrovski S., Goldstein D.B. (2016) 'Unequal representation of genetic variation across ancestry groups creates healthcare inequality in the application of precision medicine'. *Genome Biology*, 17 (1) 15 <https://doi.org/10.1186/s13059-016-1016-y>

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W. W., Hegde, M., Lyon, E., Spector, E., Voelkerding, K., & Rehm, H. L. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genetics in Medicine : Official Journal of the American College of Medical Genetics*, 17(5), 405–424.
<https://doi.org/10.1038/gim.2015.30>

Robinson P.N., Köhler S., Bauer S., Seelow D., Hom D., Mundlos S. (2008). The Human Phenotype Ontology: a tool for annotating and analysing Human Hereditary Disease. *American Journal Human Genetics*, 85 (5) 610-615 doi: 10.1016/j.ajhg.2008.09.017

Schultz K.F., Altman, D.G, Moher, D. (2010). CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *BMJ* 340. C332
<https://doi.org/10.1136/bmj.c332>

Shukralla A., Tudur-Smith C., Powell G.A., Williamson P.R., Marson A.M. (2011) Reporting of adverse events in randomised controlled trials of antiepileptic

drugs using the CONSORT criteria for reporting harms. *Epilepsy Research* 97 (1-2) 20-29 <https://doi.org/10.1016/j.eplepsyres.2011.06.015>

Trost B., Walker S., Haider S.A., Sung W.L., Pereira S., Phillips C.L. ... Scherer S.W. Impact of DNA source on genetic variant detection from human whole-genome sequencing data. *Journal of Medical Genetics* 56 (12) 809-817
<http://dx.doi.org/10.1136/jmedgenet-2019-106281>

Tukey J. (1949) Comparing Individual Means in the Analysis of Variance. *Biometrics* 5 (2) 99-114 <https://doi.org/10.2307/3001913>

Zhu Q., Hu Q., Sheperd L., Wang J. Wei L., Morrison C. ... Yao S. (2015). The impact of DNA input amount and DNA source on the performance of whole-exome sequencing in cancer epidemiology. *Cancer Epidemiol Biomarkers Prev*, 24, (8) 1207-1213 doi: 10.1158/1055-9965.EPI-15-0205

Table 1 List of thirty-eight items selected for inclusion. Items were selected based on consensus from neurologists, geneticists and bioinformatic specialists.

Section of paper	Descriptor of checklist item	Comments
Title	1) The title should state the gene or key gene of interest	Gene should be easily identified in the title
Abstract	2) Abstract should mention the key variant in the text	Variants should be clearly stated so readers can easily identify the key results
	3) Abstract should mention the gene or genes in the text	Gene should be easily identified in the abstract
	4) Abstract should state the clinical syndrome or its clinical features	Syndrome or its clinical features should be clearly stated in the abstract
Description of case or Introduction	5) Mention the clinical features of the proband, relatives or siblings in text of article or in a table	Clinical features should be clearly described
	6) Mention the ethnicity of the proband or cases	Variants may be specific to certain ethnicities. Reporting of ethnicity allow for future analysis to occur
	7) State the Human Phenotype Ontology (HPO) Number when describing the clinical features of the participants	HPO number is a standardised way of communicating phenotype, comparative analyses would be helped by standardisation
	8) State the MIM or OMIM number of clinical syndromes	OMIM numbers are relevant when stated and helps compare studies. OMIM is the most common disease database and is widely quoted in the medical literature.
	9) State the consanguinity of parents of the proband	Consanguinity is relevant when one thinks a mutation is recessively inherited
	10) If cohort study, mention how patients were included and selected	Cohort study item. Reporting of patient selection minimises bias in cohort studies
	11) Details of ancillary investigations	
	12) Details of prior genetic tests like gene panels or chromosomal studies prior to considering WES	

	13) Provide a pedigree chart	
	14) Provide a molecular science or genetics background in the introduction	Studies should report any relevant background which may be relevant for further analysis
Methods	15) State DNA source	Source of DNA impacts quality of DNA obtained and prevents miss-mapping. Helps identify germline or somatic mutations of both
	16) State exome target capture kit	Kit type can indicate quality of overall sequencing due to inexact chemistry of kits available
	17) Mention the type of sequencer used	Type of sequencer has effect on consistency and result.
	18) Mention some form of quality control of raw data from sequencer (Phred, or CG% or CG distribution)	
	19) State reference genome and build used to assemble the exome data	Variant coordinates will differ based on the type and build of reference genome.
	20) State type of Haplotype caller/bioinformatics pipeline used to distinguish variants from the reference genome	Different variant callers can produce different results; some are more accurate than others.
	21) State variant filtering approach	Were variants filtered by pre-defined thresholds; probability estimates, in reference to curated variant datasets
	22) State type of annotation databases consulted	Examples of annotation databases are ANNOVAR etc.
	23) Stat what minor allele frequency (MAF) databases were consulted	MAF is important in determining if a variant is rare or not
	24) State <i>in-silico</i> Prediction databases consulted	Damage predicting databases such as PolyPhen SIFT, CADD, or MutationTaster etc.
	25) State evolution conservation databases consulted	Evolution conservation databases inform us what parts of the gene are under constraint or drift. Genetic data under constraint almost always translates to an important part of the polypeptide

	26) State how variant impacts translated polypeptide sequence and protein function	
	27) State pathogenic variant databases consulted for previous reports of variant (Clinvar OMIM or as a narrative description)	
	28 State whether ACMG guidelines were applied when relevant	
Results	29) If cohort study must provide a diagnostic yield as a proportion or percentage	Cohort Study item
	30) If results were validated by Sanger sequencing	Sanger sequencing is a method to validate a variant, older studies would need validation but more recently it is not needed
	31) Was full transcript information of the variant was provided	Full transcript of the variant allows other readers to compare and verify the findings
	32) Do they report variants of undetermined significance	Poor reporting of VUS may indicate reporting bias
	33) State missense, insertion deletion substitution stop-gain or any other form of mutation	
	34) State variant segregation/inheritance or did it arise de novo	
	35) Do they report findings on sequence chronogram	
Discussion	36) Discuss any previous literature in view of findings	
	37) Do authors deposit their findings to ClinVar NCBI SRA or any relevant data bank	Do authors state that they contribute the results of the study to a data banks
	38) Is the patient's phenotype compared to the reported associated disease state of the disrupted gene	

Table 2 Characteristics of included studies.

Covariate		Number of studies (%) total=451
Syndrome	Developmental Epileptic encephalopathy	149 (33%)
	Epilepsy and Intellectual Disability	110 (24%)
	Other	192 (43%)
Journal type	Genetics	225 (50%)
	Other	57 (13%)
	Neurology	169 (37%)
Study type	Case/Case series	360 (80%)
	Cohort	91 (20%)
Publication date	Pre-2017	218 (48%)
	Post-2016	233 (52%)
Impact factor	Low Impact factor	188 (42%)
	High Impact factor	57 (13%)

Table 3 Top ten journals publishing WES studies.

Journal Name	Number of Studies	Percentage
American Journal of Medical Genetics	33	7.3
American Journal of Human Genetics	30	6.7
Epilepsia	27	6.0
Brain and Development	21	4.7
European Journal of Human Genetics	18	4.0
Clinical Genetics	17	3.8
Journal of Medical Genetics	14	3.1
Neurology	14	3.1
Annals of Neurology	13	2.9
Pediatric Neurology	13	2.9

Table 4 Means and difference of means.

Covariate	Number of Studies	Mean Score	Range of Scores	Difference of Means	95% CI of mean difference	p-value
Pre-2017	218	20.9	4-30	0.5	-1.4 to 0.4	0.24
Post-2016	233	21.4	8-31			

Case study	360	21.17	8-31	0.09	-1.0 to 1.18	0.86
Cohort Study	91	21.08	4-30			
Genetics Journal	225	22.4	10-30	2.5	1.6 to 3.4	<0.01
Neurology Journal	169	19.9	9-31			
Genetics Journal	225	22.4	10-30	2.3	1 to 3.5	0.001
Other Journal	57	20.1	4-30			
Neurology Journal	169	19.9	9-31	0.2	-1.7 to 1.3	0.77
Other Journal	57	20.1	4-30			
Low Impact factor	188	20.6	4-31	1.4	-2.7 to -0.05	0.04
High Impact factor	76	22.0	11-30			

Table 6 Percentages of items reported in WES articles.

	Item	Percentage
Reported well	Title mentions gene	87
	Abstract mentions gene	97
	Abstract mentions syndrome	96
	Clinical features	95
	Inclusion and selection of patients	80
	Ancillary investigations	88
	Mutation type	95
	Previous literature	98
	Phenotype compared to reported associated disease	86
	Variant segregation/inheritance	91
	Sanger sequencing Validation	77
	Molecular science background	78
	Sequencer used	73
	Minor Allele Frequency	71
Reported with medium frequency	Exome target kit	67
	How variants impacts polypeptide	64
	Abstract mentions variant	63
	Ethnicity	61
	<i>In-silico</i> predicting databases such as PolyPhen, SIFT etc.	59
	Consanguinity	58
	Haplotype caller	58
	Pedigree chart	53
	Variant filtering	48
	DNA source	47
	Pathogenic databases of previous reports	47
	Full transcript information	47
	Reference Genome used to assemble exome data	45
	Sequence chronogram	45
	Diagnostic yield estimate	43
	Annotation databases	38
	Prior Genetic tests	35
	OMIM	35

	Evolutionary conserved databases	31
Reported Poorly	VUS reported	19
	Quality scores such as Phred, CG% or CG distribution	14
	ACMG guidelines applied	13
	Deposit findings to <i>ClinVar</i>	6
	Human Phenotype Ontology (HPO) terms	3

Figure 1 Selection of Studies

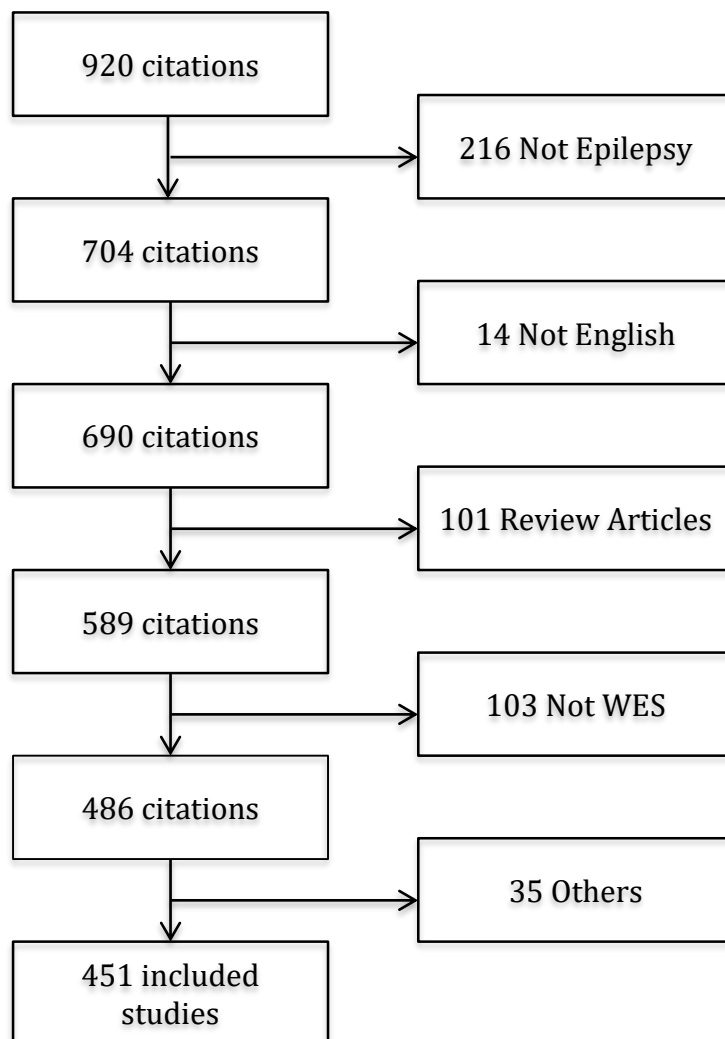


Figure 2 Forest Plot comparing Other Journals and Genetics Journals

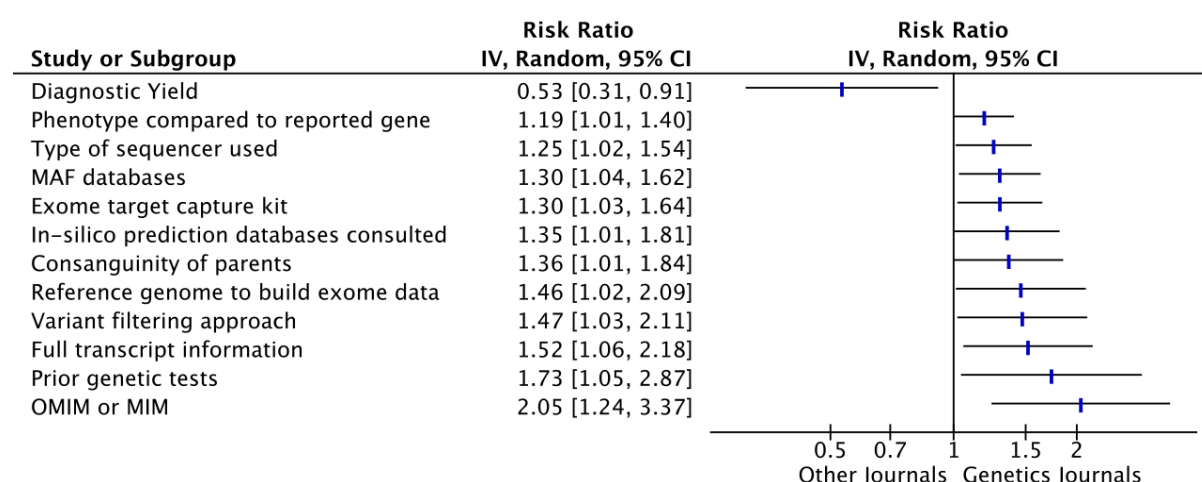


Figure 3 Forest Plot comparing Other Journals and Neurology Journals

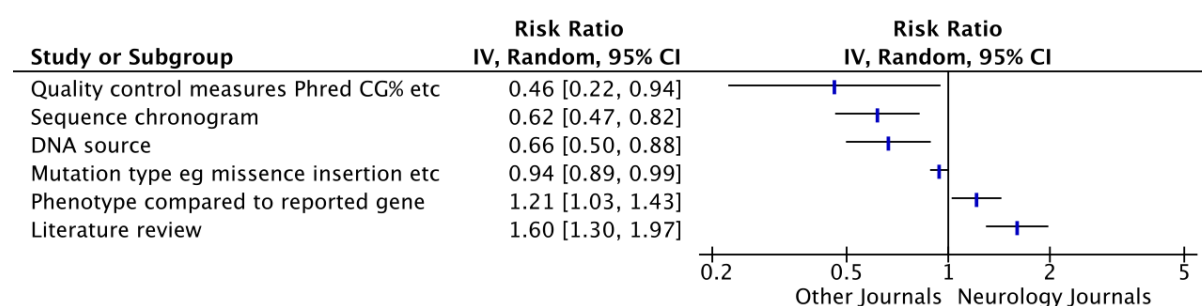


Figure 4 Forest Plot comparing Pre 2017 vs Post 2016 studies

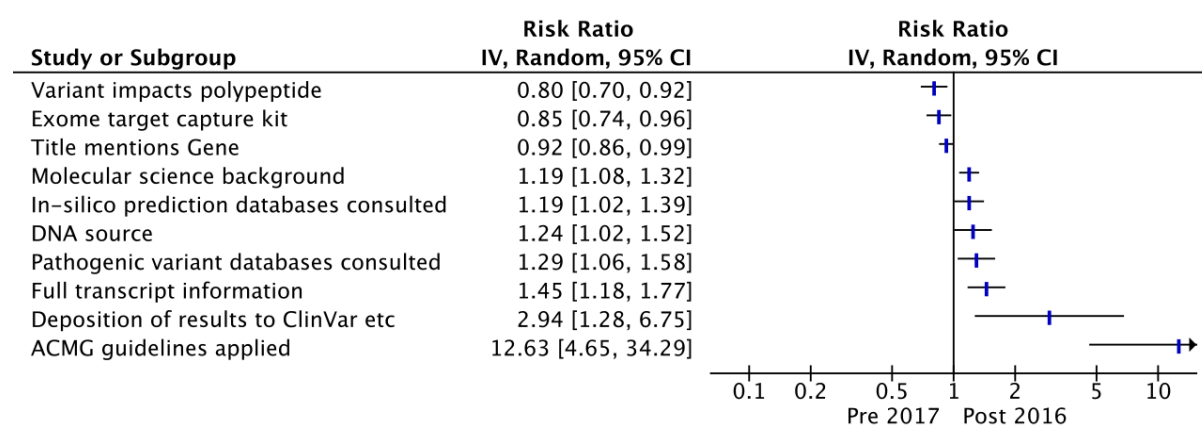


Figure 5 Forest plot comparing Cohort and Case studies

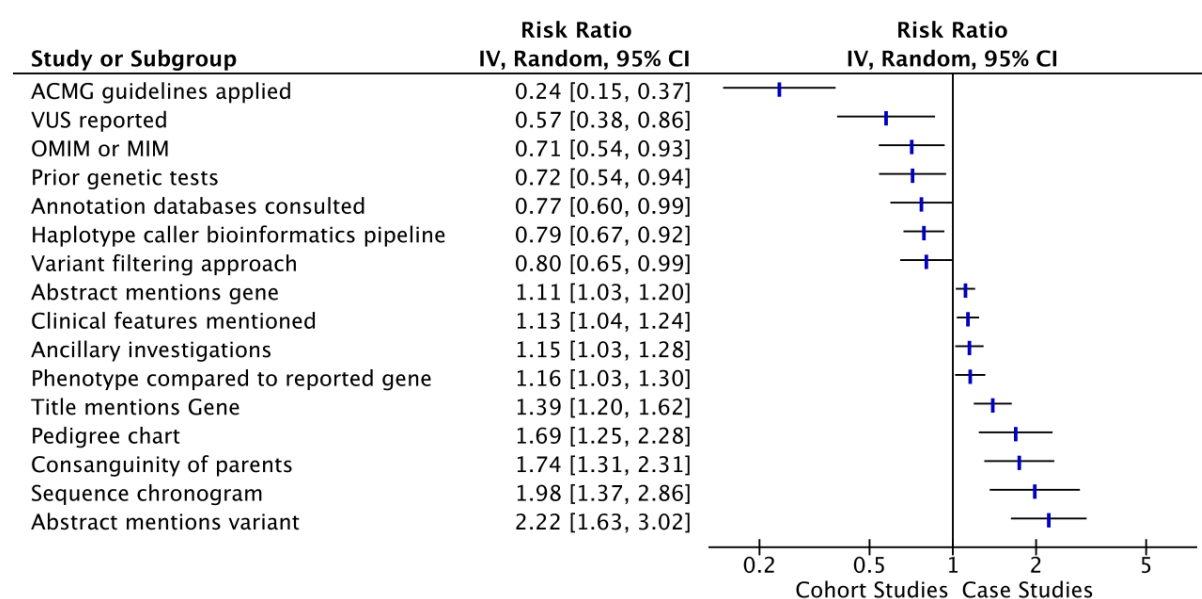


Figure 6 Forest Plot comparing High Impact versus Low Impact Journals

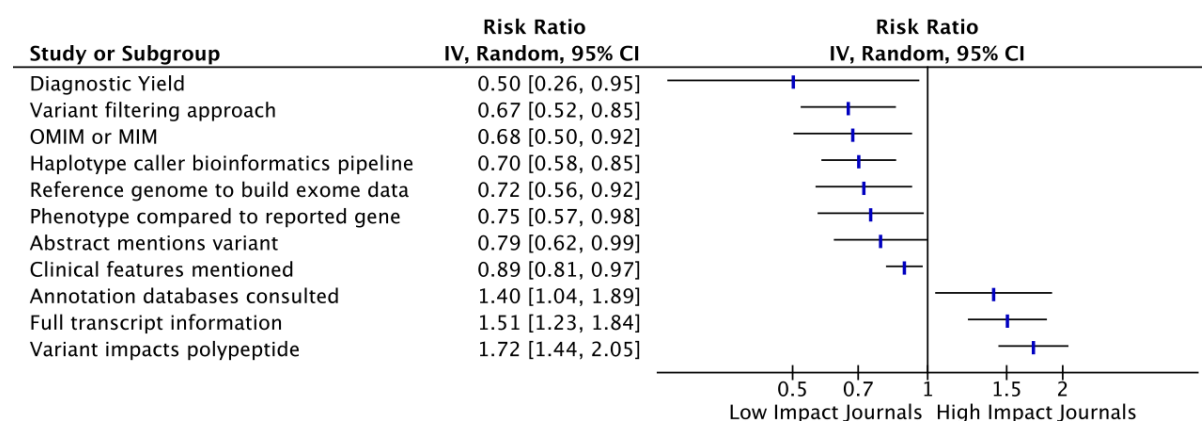


Figure 7 Forest Plot comparing Neurology and Genetics Journals

