



## OPEN The plausibility of claimed induced seismicity

Max Wilkinson<sup>1,6</sup>, Najwa Mhana<sup>1,2</sup>, Miles P. Wilson<sup>1,3</sup>, Gillian R. Foulger<sup>1,3</sup>✉, Timur Tezel<sup>4</sup> & Jon G. Gluyas<sup>3,5</sup>

Claims of industrially induced seismicity vary from indisputable to unpersuasive and yet the veracity of industrial induction is vital for regulatory and operational practice. Assessment schemes have been developed in response to this need. We report here an initial assessment of the reliability of all globally known cases of proposed human-induced earthquakes and invite specialists on particular cases to refine these results. 1235 cases were assessed, requiring over 1000 h of work. From the 881 cases for which scorable evidence is available, we class 87% as 'Confidently Induced', 10% as 'Probably Induced', 2% as 'Equivocal' and < 1% as 'Confidently Natural'. The most seismogenic activities are fracking, research, geothermal, water reservoir impoundment, conventional oil and gas. Least seismogenic activities are construction, deep penetrating bombs, coal bed methane. 354 cases (29%) lack enough information to be assessable. Future work could include applying data mining techniques including natural language processing and AI to uncover new evidence. Future best practice for rapid assessment of cases would ideally involve an independent panel of scientists who rapidly apply a questionnaire scheme, reach consensus, and inform a response.

### Main

Human-induced earthquakes are a global phenomenon caused by various activities<sup>1–7</sup>. Their negative effects vary from societal nuisance<sup>8</sup> to major economic losses and human fatalities<sup>9</sup>. There are major challenges in predicting the occurrence and magnitude of induced earthquakes. It is, however, very likely that they will become more commonplace as human populations grow and Earth's resources are increasingly exploited. To date, over 1200 scientifically proposed cases of induced seismicity have been documented globally<sup>10</sup>. The scientific evidence, and thus the reliability of claims for induced earthquakes, is highly variable in completeness and quality. Temporal trends and bias exist whereby interest in such earthquakes is linked to the proliferation of causative activities and their associated controversies. Global interest, often in terms of opposition, to fracking of shales and other low-permeability petroleum reservoirs expanded from around 2012 and increased public awareness of induced earthquakes<sup>11</sup>.

In the absence of universally applicable quantitative approaches to determine case reliability, questionnaire schemes have been developed for specific seismogenic activities<sup>12</sup>. These have evolved in complexity and applicability<sup>13–15</sup>. The most recent such scheme, "Evaluating Proposals of Induced Earthquakes (*E-PIE*)", was developed to be universally applicable regardless of causative activity. *E-PIE* is objective, based entirely on the strength of claimed evidence, and de-emphasises the personal opinions of the assessor<sup>16</sup>. Such schemes are becoming critical tools for industries where operators and regulators recognise the potential for induced seismicity and need to act rapidly if it occurs.

In this study we provide an initial, comprehensive, standardised evidence assessment of all currently known, worldwide cases of induced seismicity. A single assessor independently applied the *E-PIE* scheme to 1235 proposed cases in the Human-Induced Earthquake Database (*HiQuake*). This publicly available database is the largest and most complete compilation of scientifically proposed human-induced earthquakes<sup>17–19</sup>. The assessment took over 1000 h of study. A preferred approach would have been to use an expert panel to assess all 1235 cases, but this would have required excessive resources. The application of a consistent scheme by a single assessor provides a uniform, preliminary set of results suitable for assessing induced seismicity through time and across different industries.

Compromise was required in applying a standardised assessment scheme to the broadest range of cases and the contentious nature of such work. The results should thus be viewed as initial and provisional. We invite

<sup>1</sup>Foulger Consulting Ltd., Durham, UK. <sup>2</sup>Damascus University, Damascus, Syria. <sup>3</sup>Department of Earth Sciences, Science Labs, Durham University, Durham, UK. <sup>4</sup>Department of Earth Sciences, Science Labs, Sakarya University, Sakarya, Türkiye. <sup>5</sup>Geoenergy Durham Ltd., Durham, UK. <sup>6</sup>School of Engineering, Newcastle University, Newcastle upon Tyne, UK. ✉email: g.r.foulger@durham.ac.uk

feedback from experts on individual cases so these results, which we provide publicly in the *HiQuake* database, may be improved.

## Results

### Scoring distribution by question

We analysed the distribution of assessor responses (Fig. 1) to each of the nine questions of the *E-PIE* scheme (Fig. 2). *E-PIE* comprises nine generalized questions, responses to which indicate the likelihood that the earthquakes were induced by human activity<sup>16</sup>. Of the 1235 cases studied, the assessor found there was insufficient evidence to score *any* question for 354 cases (29%). These were scored as ‘no evidence’ across all nine questions (Fig. 1, grey dotted area). Such cases are commonly presented in tables without supporting data, a common situation for water reservoir impoundment<sup>3,20</sup>, mining<sup>21,22</sup>, conventional oil and gas<sup>5</sup> and nuclear explosions<sup>23</sup>. Questions 1–4, which refer to the temporal and spatial distributions of the seismicity in relation to the activity, were dominated by induced evidence (45–66%). Questions 6–9, which refer to auxiliary seismic parameters such as focal mechanisms, seismic swarm evolution and surface deformation, were most lacking of evidence (40–65%). Question 5, which relates to pre-industrial seismicity, had the greatest proportion of scores that were ‘equivocal’ (i.e. regional seismicity occurred before the activity, 33%) and ‘natural’ (i.e. local seismicity occurred before the activity, 12%).

### Cluster analysis

Summary scores for each case were calculated using an assessment tool developed to streamline the workflow (Online Methods: *Application of E-PIE scheme*). We found a continuum of results ranging from ‘confidently natural earthquakes’ (-1 score), through equivocal (0 score), to ‘confidently induced earthquakes’ (+1 score)<sup>16</sup>. For classification, we looked for potential clustering that might enable evidence-based division of the results into likely induced cases and likely natural cases (Online Methods: *Cluster analysis*). Excluding the ‘no evidence’ cases, we found a 4-cluster optimisation based on a Goodness of Variance Fit (GVF) of 0.93 to be the minimum possible number of clusters to represent the data (Figure S1a, triangular data point, table S2). We name these Cluster 1: Confidently Natural; Cluster 2: Equivocal; Cluster 3: Probably Induced; Cluster 4: Confidently Induced. Their bounds are shown in the frequency/cumulative-frequency plot of Fig. 3.

### Scoring distribution by process

All *HiQuake* cases are linked to one of 16 activities<sup>10</sup> (Table 1). We explored the *E-PIE* scores and statistics for each of these activities (Fig. 4; Online Methods: *Application of E-PIE scheme*). There was significant overlap in the range of *E-PIE* scores for cases within most activities. However, there were differences between activities. Figure 4 shows activities in order of reducing median *E-PIE* score. Cases in fracking, research, oil and gas/waste fluid injection, and groundwater extraction mostly lay in the Confidently Induced range with fracking being most numerous. Cases in waste fluid disposal, geothermal, and carbon capture and storage (CCS) activities generally spanned the range from Confidently Induced to Probably Induced. Cases in oil and gas, chemical explosions, mining, water reservoir impoundment, conventional oil and gas, and nuclear explosions activities were mostly within the Probably Induced range. Construction lies on the boundary between Probably Induced and Equivocal, whilst deep penetrating bombs was mostly within the Equivocal range. Coal bed methane (CBM) comprises a single case entirely within the natural range. Nuclear explosions contained the highest-scored case (the Cannikin test: 1.00). The lowest scored case is the Center, Texas, waste fluid disposal case where the earthquake sequence began before the industrial activity. In this situation the *E-PIE* assessment stopped at Question 1 because it triggered an exit criterion, resulting in the case being scored as natural (*E-PIE* score - 1.00).

The activities chemical explosions, construction, and coal bed methane had small interquartile ranges due to small case numbers. Fracking was notable for its narrow interquartile range and large number of cases. Deep penetrating bombs, geothermal, water reservoir impoundment, and conventional oil and gas had the widest interquartile ranges.

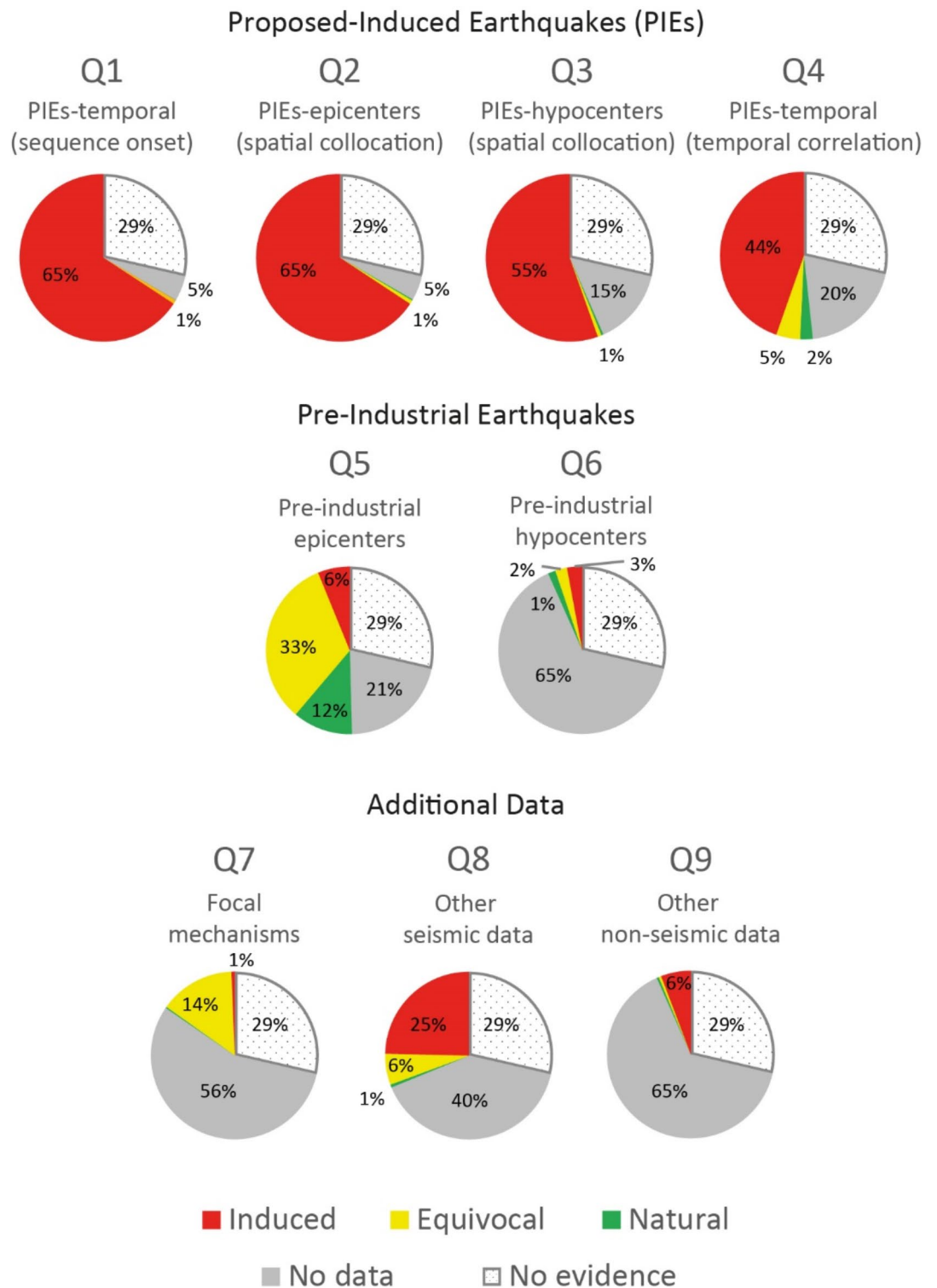
### Logical conflicts

Logical conflicts within the *E-PIE* assessment occur when there are contradictory responses to individual Questions. They indicate inconsistent evidence or assessor error and therefore act to quality check the assessment. We validated the assessment of each case using conditional logic to identify logical conflicts (see Online Methods: *Logical conflicts*). Of the 1235 cases assessed, only one logical conflict was detected – the fracking case of Fox Creek SS14. The conflict existed between a ‘natural’ score for Question 7 and an ‘induced’ score for Question 4 (Fig. 2). Reassessment of the case changed the Question 7 score to ‘equivocal’, resolving the conflict.

## Discussion

### Scientifically proposed cases are mostly reasonable

Of the 881 cases in *HiQuake* presenting scorable evidence using *E-PIE*, 87% (766) fall within the category ‘Confidently Induced’, with a range of *E-PIE* scores of 0.77–1.00 (Fig. 3, Table S2). Most cases in *HiQuake* therefore represent reasonable proposals. The remaining lie in a broad spectrum with three discrete clusters (Fig. 3, Table S2): 10% (92) lie in the ‘Probably Induced’ category, 2% (19) in the ‘Equivocal’ category, and <1% (4) in the ‘Confidently Natural’ category. ‘Confidently Induced’ cases were present for all activities except chemical explosions, construction, deep penetrating bombs, and coal bed methane (Fig. 4). The most seismogenic activities are fracking, research, geothermal (i.e. stimulation, production, injection and circulation), water reservoir impoundment, and conventional oil and gas.



**Fig. 1.** Distribution of scores for each of the nine *E-PIE* Questions for 1235 cases in *HiQuake* conducted by the single assessor. Dotted fields indicate ‘no evidence’ scores where the assessor was unable to answer any of the nine *E-PIE* Questions.

#### Key questions for determining induced or natural cause

The majority of evidence presented in favour of induced seismicity relates to *E-PIE* questions 1–4 (Figs. 1 and 2) which address the temporal and spatial nature of earthquakes relative to the industrial activity. These four questions are the most influential in determining whether the earthquakes are induced<sup>16</sup>. To reflect this importance in *E-PIE*, these questions utilise an exit criterion in question 1 and increased weighting in questions

E-PIE Assessment Scheme v.04/09/22

Assessor: NM Priority: SU Earthquake Cause: Fracking Project Name: Ryser well, Harrison County, Ohio Date: 29/10/2023

ID: 70105 Mag: 2.2 MW Mmax Date: 2013 (October) Country: USA Refs: Friberg et al. (2014), Kozłowska et al. (2018)

**Orientation**

How plausible is the proposed induction mechanism?

- Is it a well-established phenomenon?
- Is it reported for multiple localities?

How wide, in space and time, is the range of likely environmental modulation?

- Only near-field, rapid response to operations likely
- Up to medium-field, medium-term response to operations likely
- Out to far-field, delayed response to operations likely

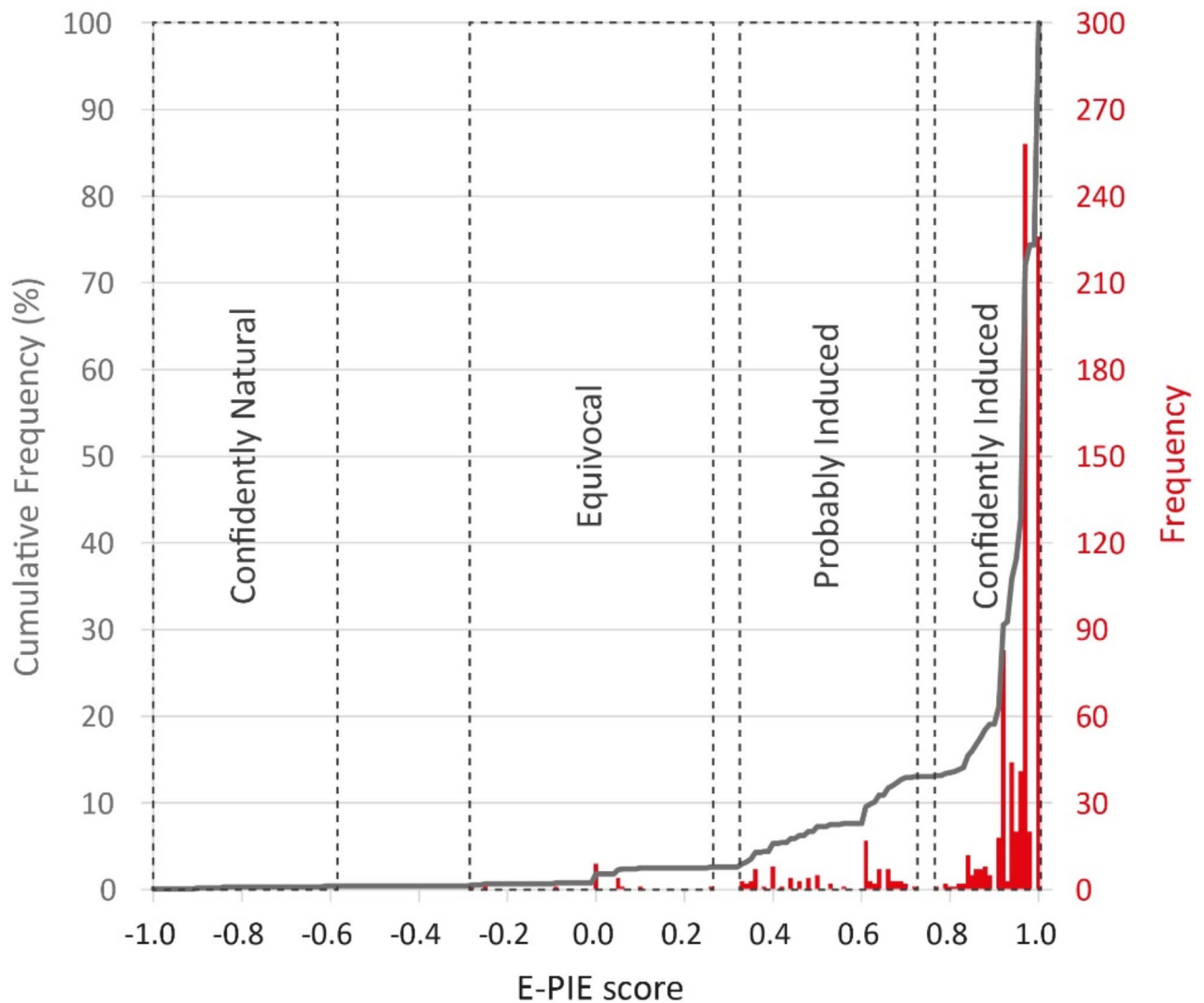
Proposed-induced earthquakes (PIEs)	Supporting Evidence
<b>(1) PIEs-temporal: Did the PIE sequence onset before, during or after the industrial activity?</b> 10	Paste Q1 evidence
(a) Insufficient information available	<input type="radio"/>
(b) The PIE sequence began before the onset of the industrial activity	<input checked="" type="radio"/>
(c) The PIE sequence began while the industrial activity was minimal OR after its cessation	<input type="radio"/>
(d) The PIE sequence began while the industrial activity was substantial	<input type="radio"/>
<b>(2) PIEs-epicenters: Is there spatial collocation between the PIEs and the likely area of environmental modulation by the industrial activity?</b> 100	Paste Q2 evidence
(a) Insufficient information available	<input type="radio"/>
(b) The PIEs are outside the likely area of environmental modulation by the industrial activity	<input checked="" type="radio"/>
(c) The PIEs are peripheral to the likely area of environmental modulation by the industrial activity	<input type="radio"/>
(d) The PIEs are within the likely area of environmental modulation by the industrial activity	<input type="radio"/>
<b>(3) PIEs-hypocenters: Is there spatial collocation between the PIEs and the likely volume of environmental modulation by the industrial activity?</b> 100	Paste Q3 evidence
(a) Insufficient information available	<input type="radio"/>
(b) The PIEs are beneath the likely volume of environmental modulation by the industrial activity	<input checked="" type="radio"/>
(c) The PIEs are peripheral to the base of the likely volume of environmental modulation by the industrial activity	<input type="radio"/>
(d) The PIEs are within the likely volume of environmental modulation by the industrial activity	<input type="radio"/>
<b>(4) PIEs-temporal: Is there temporal correlation between the PIEs and specific industrial events?</b> 100	Paste Q4 evidence
(a) Insufficient information available	<input type="radio"/>
(b) There is little or no temporal correlation between the PIEs and specific industrial events	<input checked="" type="radio"/>
(c) There is weak temporal correlation between the PIEs and specific industrial events	<input type="radio"/>
(d) There is strong temporal correlation between the PIEs and specific industrial events	<input type="radio"/>
<b>Pre-industrial earthquakes</b>	
<b>(5) Pre-industrial earthquakes-epicenters: Is there evidence for pre-industrial earthquakes at or near the site of the PIEs?</b> 10	Paste Q5 evidence
(a) Insufficient information available	<input type="radio"/>
(b) Pre-industrial earthquakes occurred at or near the site of the PIEs	<input checked="" type="radio"/>
(c) Pre-industrial earthquakes occurred in the wider region around the site of the PIEs	<input type="radio"/>
(d) Pre-industrial earthquakes did not occur at or near the site of the PIEs or in the wider region around it	<input type="radio"/>
<b>(6) Pre-industrial earthquakes-hypocenters: Is there evidence for pre-industrial earthquakes in the same volume as the PIEs?</b> 10	Paste Q6 evidence
(a) Insufficient information available	<input type="radio"/>
(b) Pre-industrial earthquakes occurred at or near the site of the PIEs at similar or shallower depths	<input checked="" type="radio"/>
(c) Pre-industrial earthquakes occurred in the wider region around the site of the PIEs at similar or shallower depths	<input type="radio"/>
(d) Pre-industrial earthquakes did not occur at or near the site of the PIEs or in the wider region around it at similar or shallower depths	<input type="radio"/>
<b>Additional data</b>	
<b>(7) Focal mechanisms: Are the focal mechanisms consistent with a natural and/or induced earthquake cause?</b> 10	Paste Q7 evidence
(a) Insufficient information available	<input type="radio"/>
(b) The focal mechanisms ARE consistent with the regional stress and NOT consistent with the proposed induction mechanism	<input checked="" type="radio"/>
(c) The focal mechanisms ARE consistent with the regional stress and ARE consistent with the proposed induction mechanism OR The focal mechanisms are NOT consistent with the regional stress and NOT consistent with the proposed induction mechanism	<input type="radio"/>
(d) The focal mechanisms are NOT consistent with the regional stress and ARE consistent with the proposed induction mechanism	<input type="radio"/>
<b>(8) Other-seismic data: Are there other seismic data to support a natural or induced cause, e.g., swarm, foreshock-aftershock pattern, b-value, total number of earthquakes, stress release corresponding to the earthquake magnitude or seismicity?</b> 10	Paste Q8 evidence
(a) Insufficient information available	<input type="radio"/>
(b) Other seismic data support a natural origin	<input checked="" type="radio"/>
(c) Other seismic data are equivocal	<input type="radio"/>
(d) Other seismic data support an induced origin	<input type="radio"/>
<b>(9) Other-non-seismic data: Are there non-seismic data that support a natural or induced cause, e.g., direct nucleation effects, precursory surface deformation?</b> 10	Paste Q9 evidence
(a) Insufficient information available	<input type="radio"/>
(b) The non-seismic data support a natural origin	<input checked="" type="radio"/>
(c) The non-seismic data are equivocal	<input type="radio"/>
(d) The non-seismic data support an induced origin	<input type="radio"/>

Submit Submit: Insufficient evidence

**Fig. 2.** The interactive form used in the *E-PIE* assessment.

2–4. Questions 5 and 6 examine whether the area was seismically active prior to the implicated activity (Fig. 2). If so, this weakens the case that earthquakes were induced (Fig. 1). These questions are least prominent in the evaluation of mining and water reservoir impoundment cases where pre-existing seismicity is commonly not considered or presented. Questions 7–9 provide the opportunity to add evidence from additional data (Fig. 2). These data may comprise analyses not yet developed, thus ensuring *E-PIE* remains applicable if new methods are developed. These questions also enable inclusion of diverse information from a range of activities. Well-documented activities including fracking, geothermal, groundwater extraction, nuclear explosions, oil and gas,





**Fig. 3.** Visualisation of the four clusters overlain on frequency and cumulative frequency plots for the 881 cases in *HiQuake* which could be scored.

research, waste fluid disposal, and water reservoir impoundment commonly provide evidence that can be input into these questions.

*E-PIE* scores vary widely within each activity class (Fig. 4). The median score and interquartile range for each activity distinguishes the least-seismogenic (construction, deep penetrating bombs and coal bed methane) from more-seismogenic activities. The single coal bed methane case is the only one scored as natural because of the earthquakes' distance and the induced stress field being incompatible with the focal mechanism<sup>24</sup>. There is significant overlap in the results for the remaining activity classes such that the *E-PIE* score of any individual case is not a reliable indicator of the claimed activity. This implies that for induced seismicity in areas of mixed activity, the assessment results cannot be used to determine the causative activity.

The ranking of activities based on their median score (Fig. 4) provides an indication of the likelihood seismicity was induced when considering the evidence collated within *HiQuake* and the variable numbers of individual cases within each activity. The interquartile range for each activity provides insight into scoring consistency during this initial assessment. Considering median and interquartile range in this context, fracking (409 cases), oil and gas/waste fluid injection (4 cases), and research (14 cases) are scored as Confidently Induced with relatively narrow ranges due to consistently comprehensive documentation of evidence.

#### Utilisation of *E-PIE* and other questionnaire schemes

Major advantages of questionnaire schemes such as *E-PIE* are their simplicity and rapid application. Whereas detailed scientific studies might provide more robust assessments, rarely are these possible within the hours to days required by operators or regulatory bodies during activities. This problem was highlighted by the 2022  $M_w$  5.2 Peace River earthquake, Alberta, Canada. The Alberta Energy Regulator initially stated the earthquake was

Activity Class	Impact on shallow crust
CCS	Pore-pressure increase
Chemical explosion	Inelastic radial deformation to form explosion cavity, subsequent collapse
Coal Bed Methane (CBM)	Pore-pressure decrease (extraction)
Construction	Vertical loading and compaction, pore-pressure increase by fluid migration
Conventional Oil and Gas	Pore-pressure decrease
Deep penetrating bombs	Inelastic radial deformation to form explosion cavity, subsequent collapse
Fracking	Pore-pressure increase
Geothermal	Pore-pressure increase (injection), pore-pressure decrease (production), thermal changes
Groundwater extraction	Pore-pressure decrease (extraction)
Mining	Vertical unloading, pore-pressure decrease, cavity collapse
Nuclear explosions	Inelastic radial deformation to form explosion cavity, subsequent collapse
Oil and Gas	Pore-pressure increase (injection)
Oil and Gas/Waste fluid injection	Pore-pressure increase (injection)
Research	Pore-pressure increase (injection), pore-pressure cycling (stimulation)
Waste fluid disposal	Pore-pressure increase (injection)
Water reservoir impoundment	Vertical loading and compaction, pore-pressure increase by fluid migration

**Table 1.** The 16 activity classes in *HiQuake* and their impact on the shallow crust.

natural<sup>25</sup> but later conclusively demonstrated it was induced by wastewater disposal. The latter conclusion was also reached retrospectively by an expert panel applying both the Verdon<sup>15</sup> and *E-PIE* schemes<sup>26</sup>.

Rapid assessment holds value in sometimes producing unexpected results that may prompt further examination and discussion. Best practice for rapid assessment of potentially induced earthquakes may comprise a readily available independent panel of scientists across regulation, industry, and academia. Panel members may apply the same questionnaire scheme and consensus may inform the course of action. A statistical analysis of the variation in questionnaire results between individual scientists and schemes, and the effect of averaging results, has been presented<sup>16</sup>. An independent panel could also debate historic cases of induced seismicity. Repeat assessments of cases by the same, or new panels, could be carried out as needed to consider new observations, data, and/or analytical methods.

### Cases lacking published evidence

A significant proportion (29%) of cases in *HiQuake* lack supporting evidence, despite being proposed in scientific literature, and thus could not be assessed using *E-PIE*. These cases are commonly listed in published tables of known induced seismicity and hence included in *HiQuake*. They are prevalent in the activities water reservoir impoundment<sup>3,20</sup>, mining<sup>21,22</sup>, conventional oil and gas<sup>5</sup> and nuclear explosions<sup>23</sup>. Cases may be based on general opinion, or the supporting evidence may lie in un-referenced or unpublished material. We made every effort to include all material available during the assessment process. However, significant potential clearly exists to extend the results to these cases by applying data mining techniques such as natural language processing and artificial intelligence to uncover new or overlooked evidence in the vast literature that now exists.

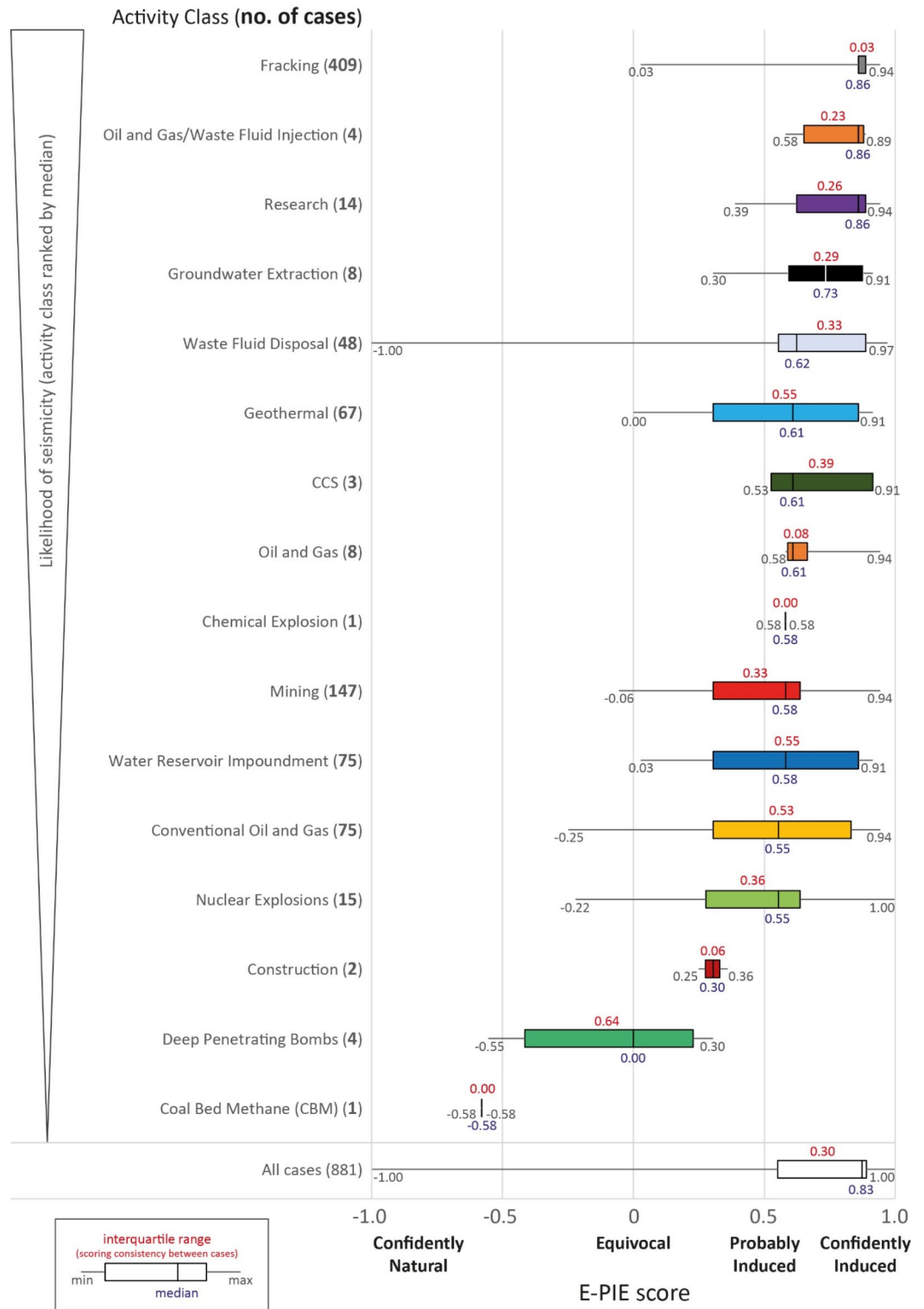
## Online methods

### Application of the E-PIE scheme

The *E-PIE* questionnaire scheme<sup>16</sup> (Fig. 2) was applied to the entire *HiQuake* database of proposed human-induced earthquakes. *HiQuake* contained 1235 cases as of 10th December 2021, the version used for this study. A single assessor with a Ph.D. in earthquake seismology worked for around 1000 h over a 20-month period to perform this task. To facilitate the assessment process and data management of the results, *E-PIE* was coded into an interactive form. Using this form, each of the nine questions in *E-PIE* must be answered with one of four responses: 'no data', 'natural', 'equivocal' or 'induced'. Answers represent the evidence proposed in the scientific literature and not the assessor's opinion of this evidence. In this way, the results are as objective as possible.

The scientific literature used to assess each case was illustrative, not comprehensive, due to the vast volume that exists for many cases. For publication-rich cases only prominent and commonly cited literature (e.g. publications in leading peer-reviewed scientific journals commonly cited throughout induced seismicity research) were included. In cases where no data were found to answer any *E-PIE* question, the case was labelled 'insufficient evidence' to distinguish it from cases where individual questions are scored 'no data'. In the rare instances where proposed evidence for a question was contradictory, the assessor prioritised the proposed evidence relating to the most rigorous scientific process, including validation of results using additional methodologies. Where proposed evidence was contradictory but considered of equal credibility the assessor recorded a response of 'equivocal'.

To visualise and interpret the collective results of the *E-PIE* assessments, a single quantifiable score was calculated for each case<sup>16</sup> as follows. The following scores were applied to the response to each question: 0 for a (no data), -1 for b (natural), 0 for c (equivocal), and 1 for d (induced). The scores are then normalised by their *E-PIE* weightings for each question (Questions 1 and 5–9 are each weighted 1/36th and questions 2, 3 and 4 are each weighted 10/36th). Summing produces an aggregate score analogous to the Induced Assessment Ratio



**Fig. 4.** Number, *E-PIE* scores and summary statistics for cases within for each of the 16 activity classes in *HiQuake*. Classes are independent of one another and arranged by their median, which has not been normalised.

(IAR)<sup>15</sup>. The weighted proportion of ‘no data’ responses is totalled and subtracted from 1 to produce a value of ‘coverage’ in the range of 0 to 1. The aggregate score is then multiplied by the coverage to produce the final *E-PIE* score in the range – 1 (natural), through 0 (equivocal) to + 1 (induced).

The results were interpreted by question and activity. For each of the 16 activities, summary statistics of the minimum, maximum, first quartile, third quartile and median value were calculated.

### Cluster analysis

Calculated *E-PIE* scores of the 1235 *HiQuake* cases yield a continuum from natural (-1), through equivocal (0), to induced (+1). Clustering exists within this continuum, providing an objective way to informally group and reference cases. Jenks Natural Break Optimisation<sup>27</sup> was used to determine discrete clusters, whereby the variance within each cluster is minimised while the variance between clusters is maximised. The optimisation is expressed as the Goodness of Variance Fit (GVF), with values calculated for cluster sizes two through eight. Four clusters, with GVF of 0.93, were selected as the minimum number of clusters with sufficient GVF. This maintained most cluster break points in the six- and seven-cluster optimisations, while minimising inter-cluster spacing (Figure S1a, Table S2, Figure S5). Kernel density estimation (KDE)<sup>28</sup> was conducted for the *HiQuake* population, which supports the four-cluster optimisation (Figure S1b).

We checked for underlying clustering inherent to the *E-PIE* scheme by producing a population of all possible permutations of *E-PIE* (Figure S3a) and by taking a random subset of the 1235 scores (Figure S3b). Comparable cluster analysis<sup>27</sup> and KDE<sup>28</sup> was conducted with the random subset (Figure S4a, S4b) and the results compared to those utilising the *HiQuake* population (Figure S5). Cluster break points within the random subset and *HiQuake* populations were found not to coincide and crucially the *HiQuake* clusters spanned break points from the random subset in multiple cases across all cluster optimisations (Figure S5). This demonstrates that clustering inherent to the *E-PIE* scheme did not significantly influence the *HiQuake* population and the resultant cluster analysis.

### Logical conflicts

The nine *E-PIE* questions were answered independently by the assessor. Each question is designed to be independent and not overlap logically with other questions since this would result in some data being counting twice (with the exception of question 1 when the response ‘natural’ is selected<sup>16</sup>). The combinations of plausible logical conflicts<sup>16</sup> within *E-PIE* are shown in Table S6. ‘Natural’ responses to questions 1, 2, 3 or 7 (Fig. 2) imply a natural origin is required and ‘induced’ responses to questions 4 or 7 (Fig. 2) imply an induced origin is required. All other responses allow ambiguity in the final result and so either solution is permitted.

A conditional statement was used to search for logical conflicts within the assessor’s responses for each case. A logical conflict exists if the following conditional statement (written here in no specific code) is met:

IF (

(Q1 = ‘natural’ OR Q2 = ‘natural’ OR Q3 = ‘natural’ OR Q7 = ‘natural’)

AND

(Q4 = ‘induced’ OR Q7 = ‘induced’)

)

### Data availability

The *HiQuake* database is freely available to download from [www.inducedearthquakes.org](http://www.inducedearthquakes.org). The *E-PIE* scores are available within the supplementary data and will be included online in a future update of *HiQuake*.

Received: 22 July 2024; Accepted: 27 November 2024

Published online: 28 December 2024

### References

1. Ellsworth, W. L. Injection-induced earthquakes. *Science* **341**, 1225942 (2013).
2. Foulger, G. R., Wilson, M. P., Gluyas, J. G., Julian, B. R. & Davies, R. J. Global review of human-induced earthquakes. *Earth Sci. Rev.* **178**, 438–514 (2018).
3. Gupta, H. K. A review of recent studies of triggered earthquakes by artificial water reservoirs with special emphasis on earthquakes in koyna, India. *Earth Sci. Rev.* **58**, 279–310 (2002).
4. McGarr, A., Simpson, D. & Seeber, L. Case histories of induced and triggered seismicity. *Int. Handb. Earthq. Eng. Seismology*. **81**, 647–664 (2002).
5. Nicholson, C. & Wesson, R. L. Triggered earthquakes and deep well activities. *Pure. appl. Geophys.* **139**, 561–578 (1992).
6. Schultz, R. et al. Hydraulic fracturing-induced seismicity. *Rev. Geophys.*, **58**, e2019RG000695 (2020).
7. Wilson, M. P. et al. Anthropogenic earthquakes in the UK: A national baseline prior to shale exploitation. *Mar. Pet. Geol.* **68**, 1–17 (2015).
8. Clarke, H., Eisner, L., Styles, P. & Turner, P. Felt seismicity associated with shale gas hydraulic fracturing: The first documented example in Europe. *Geophys. Res. Lett.* **41**, 8308–8314 (2014).



9. Zhao, Y. et al. The 2021 ms 6.0 luxian (China) earthquake: Blind reverse-fault rupture in deep sedimentary formations likely induced by pressure perturbation from hydraulic fracturing. *Geophys. Res. Lett.* **50**, e2023GL103209 (2023).
10. Wilson, M. P., Foulger, G. R., Gluyas, J. G., Davies, R. J. & Julian, B. R. Hiquake: the human-induced earthquake database. *Seismol. Res. Lett.* **88**, 1560–1565 (2017).
11. Wright, M. Making it personal: How anti-fracking organisations frame their messages. *J. Politics Soc.* **24**, 105–123 (2013).
12. Davis, S. D. & Frohlich, C. Did (or will) fluid injection cause earthquakes?-criteria for a rational assessment. *Seismol. Res. Lett.* **64**, 207–224 (2013).
13. Davis, S. D., Nyffenegger, P. A. & Frohlich, C. The 9 April 1993 earthquake in south-central texas: Was it induced by fluid withdrawal? *Bull. Seismol. Soc. Am.* **85**, 1888–1895 (1995).
14. Frohlich, C. et al. A historical review of induced earthquakes in Texas. *Seismol. Res. Lett.* **87**, 1022–1038 (2016).
15. Verdon, J. P., Baptie, B. J. & Bommer, J. J. An improved framework for discriminating seismicity induced by industrial activities from natural earthquakes. *Seismol. Res. Lett.* **90**, 1592–1611 (2019).
16. Foulger, G. R. et al. Human-induced earthquakes: E-pie—a generic tool for evaluating proposals of induced earthquakes. *J. Seismolog.* **27**, 21–44 (2023).
17. Kivi, I. R. et al. Global physics-based database of injection-induced seismicity. *Earth Syst. Sci. Data Discuss.* **15**, 3163–3182 (2023).
18. Kundu, B. et al. Influence of anthropogenic groundwater pumping on the 2017 november 12, m7.3 iran-iraq border earthquake. *Geophysical Journal International*, **218**, 833–839 (2019).
19. Witze, A. Risk of human-triggered earthquakes laid out in biggest-ever database. *Nature* <https://doi.org/10.1038/nature.2017.22693> (2017).
20. Guha, S. K. & Patil, D. N. Large water-reservoir-related induced seismicity. *Gerlands Beiträge Zur Geophysik.* **99**, 265–288 (1990).
21. Li, T., Cai, M. F. & Cai, M. A review of mining-induced seismicity in China. *Int. J. Rock. Mech. Min. Sci.* **44**, 1149–1171 (2007).
22. Iannacchione, A. T. & Zelanko, J. C. Occurrence and remediation of coal mine bumps: a historical review in *proceedings of mechanics and mitigation of violent failure in coal and hard-rock mines*. Special Publication 01–95, 27–67 (U.S. Bureau of Mines 1995).
23. Hamilton, R. M., Smith, B. E., Fischer, F. G. & Papanek, P. J. Earthquakes caused by underground nuclear explosions on pahute mesa, nevada test site. *BSSA* **62**, 1319–1341 (1972).
24. Albano, M. et al. Did anthropogenic activities trigger the 3 April 2017 mw 6.5 Botswana earthquake? *Remote Sens.* **9**, 1028 (2017).
25. Schultz, R. et al. Disposal from in situ bitumen recovery induced the ml 5.6 peace river earthquake. *Geophys. Res. Lett.* **50**, e2023GL102940 (2023).
26. Salvage, R. O., Eaton, D. W., Furlong, C. M., Dettmer, J. & Pedersen, P. K. Induced or natural? Toward rapid expert assessment, with application to the mw 5.2 peace river earthquake sequence. *Seismol. Res. Lett.* **95**, 758–772 (2024).
27. Jenks, G. F. The data model concept in statistical mapping. *Int. Yearbook Cartography.* **7**, 186–190 (1967).
28. Rosenblatt, M. Remarks on some nonparametric estimates of a density function. *Ann. Math. Stat.* **27**, 832–837 (1956).

## Acknowledgements

This paper is a continuation of work originally initiated and funded by the Nederlandse Aardolie Maatschappij BV (NAM), Schepersmaat 2, 9405 TA Assen, The Netherlands.

## Author contributions

The assessor is anonymous. Max Wilkinson analysed the final results and drafted the paper and figures. Miles P. Wilson, Gillian R. Foulger and Najwa Mhana developed the paper's structure. Timur Tezel and Jon G. Gluyas contributed to the results and discussion. All authors contributed to the final manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-81632-3>.

**Correspondence** and requests for materials should be addressed to G.R.F.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024