

Improving Child Protection by integrating research evidence and clinical experience

Dr David M Foreman MB MSc FRCPsych FRCPC* (David_Foreman@doctors.net.uk)

Institute of Psychiatry & Isle of Man Govt

Prof Nicky Best PhD (nicky.x.best@gsk.com)

GlaxoSmithKline and Imperial College London

Prof Frank Dunstan PhD (DunstanFD@cardiff.ac.uk)

Cardiff University

Prof Neil McIntosh MD FRCP FRCPC (neil.mcintosh@ed.ac.uk)

University of Edinburgh

Prof Deborah Ashby PhD (deborah.ashby@imperial.ac.uk)

Imperial College London

Author contributions: DMF originated the project, proposed the original concept, drafted and edited this paper; NM chaired the committee, led the literature review and (with DF) formulated the clinical topics of this paper; NB conducted the analysis, and contributed the figures; FD and DA constructed and critiqued the analytic model with NB; all authors co-wrote the text.

*Corresponding Author

Acknowledgments

- This work was commissioned by a committee under the auspices of the Royal College of Paediatrics and Child Health (RCPCH), and including members of the Royal Statistical Society (RSS).
- Financial support for the committee's work was provided by the RCPCH, but the wider College took no part in determining the committee's work.
- All the authors were members of this committee
- No author has competing interests relevant to the topic of this paper to declare.
- This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Word count 2333

Contact details for corresponding author:

Address: 65 Derby Square Douglas IM1 3LR

Telephone: 07624 482989

Email: David_Foreman@doctors.net.uk

Abstract

We present a method to reduce the risk of psychological bias and error in judgments of child abuse. We use published research on the presenting evidence to create a credible range for the probability that the evidence is indicative of abuse (formally, a “90% credible interval” covering the most plausible 90% of values for this probability), which experts may then use to adjust their opinions about the presenting case. For illustration, we provide an example that misled experts and caused great public concern.

The method entails writing and specifying a Bayesian statistical model for the combined probability of the observed signs indicating abuse. In the illustration, the literature was systematically searched for studies associating three possible signs of child abuse – an Acute Life-Threatening Event (ALTE), oronasal bleeding, and Sudden Infant Death Syndrome (SIDS) in a sibling, whose concurrence many experts considered to strongly indicate abuse, while the Courts found otherwise. With probability expressed as a decimal between 0 (impossible) and 1 (certain), the 90% credible interval for abuse associated with a combination of ALTE and oronasal bleeding lay between 0.06 and 0.51 (i.e. there is only a 10% chance that the probability of abuse either lies below 0.06 or above 0.51, in the absence of additional relevant information). Sibling death from SIDS only increased the probability of abuse if it was assumed to occur independently of the other signs: if (as was more likely in practice) causes other than abuse could also link these signs, then this additional sign did no more than increase uncertainty, shown by a wider 90% credible interval for the probability of abuse (0.03 to 0.73). An expert would need other case-related data that specifically contradicted the research so summarised in order to give a rational opinion on the probability of abuse outside these limits. This method could thus support expert opinion in complex cases when data is sparse, by modelling the impact of different assumptions that are often made unconsciously and

implicitly in human judgments, while not replacing human judgment with an algorithm.

Keywords

Child Protection; Expert Witness; Bayesian modelling; Detection Error; Assessment; Factitious Illness and Injury; Munchausen Syndrome by Proxy

Introduction

The effective management of child protection involves good practice, good law, and effective communication about individual cases between the professional and legal systems involved. Paediatricians' roles in child protection exemplify this process well, and their difficulties mirror those of practitioners in other disciplines. Within the medical system, they need to be able to determine the probability of abuse, given presenting symptoms and signs. Should abuse be likely, part of its management then requires the paediatrician to be able to present findings to the legal system, which allows appropriate legal action to be taken to protect the child and, if appropriate, punish the perpetrator. In the courts, the paediatrician functions as an expert witness i.e., a witness whose opinions are taken as evidence by the court, additional to any factual evidence it accepts. In the United Kingdom, public confidence in medico-legal paediatric assessments was undermined by two high-profile legal cases about expert evidence (Meadow v General Medical Council 2006; Council for the Regulation of Healthcare Professionals v General Medical Council & Anor 2005). The reverberations caused UK paediatricians to become less willing to work with potential child abuse cases (Mathews et al. 2009) or make diagnoses of abuse (González-Izquierdo et al. 2010). US paediatricians' perceptions of child abuse processes also moderated their willingness to report or investigate abuse (Flaherty, Sege, and Hurley 2008), resulting in inconsistent reporting and systemic bias (Sege and Flaherty 2008; Lindberg et al. 2013). More generally, there have been reports of poor standards in expert witnesses in the UK and Canada (Ireland 2012; Cradock 2011), while concerns about the validity of expert witness testimony are reported in the US (Johnston and Sartwelle 2013). The UK has recently moved to reduce the use of any kind of expert witness in Family Court proceedings (Ministry of Justice and Department of Education 2014), though elsewhere expert witnesses' explanations have been valuable in assisting juries to appreciate the significance of counter-intuitive evidence (Seymour et al. 2014). So, there is a

continuing need to improve both the reliability and validity of expert assessments of child abuse, and the confidence any legal system may place in them.

For any practitioner, the diagnosis of child abuse is challenging, and the pressure to be accurate immense. If the history has been obtained from the perpetrators, it may be untrustworthy. Presenting signs may be suggestive of a range of conditions other than abuse. There is thus considerable potential for professional opinion to be biased by “framing effects.” These occur because there is a psychological tendency to be averse to, with, in consequence, an attempt to avoid, situations perceived as potential losses (or, equivalently, harm), irrespective of the true balance of risk and benefit involved (Tversky and Kahneman 1981). With abuse, a risk for false positive diagnosis might arise because even expert practitioners may overestimate the probability of rare but significant events (Lichtenstein et al. 1978), particularly if the suspected event would trigger an action by the practitioner that would protect a child from possible harm (Tversky and Kahneman 1981). Conversely, fear of professional censure for an erroneous diagnosis, if sufficiently intense, could lead to a framing effect towards under-detection, as found by González-Izquierdo et al. (2010). The simple collection of additional confirmatory data also has risks, as excessive information has been associated with impaired practitioner comprehension (White et al. 2010) and catastrophic false negative judgments about children being at risk of significant harm (Munro 2010).

Making use of the literature is an obvious solution to these problems, but integrating practitioners’ assessments with research-based probabilities of abuse is fraught. Research on suspected signs may be sparse or absent, while signs themselves may combine in a multitude of ways, which might not have been researched jointly, even if individual signs have been studied. Some forms of evidence, such as uniform case series, could even contribute to false positive judgments if the prevalence is too low for

the series' sampling frame to capture counter-examples, or if alternative causes to abuse have not been sought. While attempts have been made to correct for bias in the quantification of expert opinion (Kynn 2008; Johnson et al. 2010) these involve expert surveys, which are clearly not appropriate for the assessment of abuse.

The method we describe below derives a likely range for the probability of abuse (expressed as a decimal between 0 and 1, where 0 = impossible and 1 = certain) in an individual case, based on the relevant literature about its presenting signs and symptoms. For a while, this range is called a "90% credible interval", and is defined as being 90% of the full range of probabilities calculated. Interpreting a credible interval requires understanding that there is only a 10% chance of the true probability being below the bottom, or above the top end of the calculated range. Narrow credible intervals indicate that the true probability can be estimated with a high degree of confidence, while wide credible intervals suggest the reverse. The practitioner can then use this range as a set of "statistical calipers" to set reasonable bounds about the judgment of likelihood of abuse. Its particular strength is its ability to combine the evidence from many different signs, potentially taking different interactions into account, thus allowing estimation of the probability of abuse when multiple signs are present that might not have been researched together. Following a description of the method, we demonstrate its application to the expert opinions associated with both of the UK cases mentioned above (*Meadow v General Medical Council* 2006; *Council for the Regulation of Healthcare Professionals v General Medical Council & Anor* 2005). The experts' belief (*General Medical Council* 2008) was that a combination of an Acute Life-Threatening Event (ALTE) and oronasal bleeding predicted abuse, particularly if there had been a previous sibling death classified as SIDS in the family; however, the Courts disagreed.

Method

Our method uses Bayesian reasoning to estimate the probabilities of observing specific signs and symptoms in abused and non-abused children from published data, and convert these into a probability that a child presenting with these signs has indeed been abused. Box 1 explains the key features of the Bayesian theory.

Box 1: Bayesian concepts

The fundamental aim of Bayes theorem is to formalize how information about one event can give us understanding of another. A classic application of Bayes theorem, and one which we draw on closely in our method, is to diagnostic testing. Diagnostic tests provide either a positive or negative test result for an event (e.g. presence of a disease), but this is not the same as actually having the event (disease). Tests are flawed and can detect things that don't exist (false positive), and miss things that do exist (false negative). Bayes theorem calculates the actual probability of the event (disease) given that a positive or negative test result has been observed.

Bayes' theorem can be expressed mathematically as follows:

$$Odds(disease | positive test) = Odds(disease) \times \frac{Pr(positive test | disease)}{Pr(positive test | no disease)}$$

The vertical bar | denotes conditionality, and may be read as "given"; the odds of an event is just the ratio of probabilities: $Pr(\text{event occurs})/Pr(\text{event does not occur})$. It tells us how much more likely it is for the event to occur than to not occur.

Intuitively, Bayes' theorem tells us how a new piece of evidence (e.g. a positive test result) changes the odds in favour of an event (e.g. disease) having occurred. This depends on 2 things – how likely the event was before we observed the test result (the first quantity on the right hand side of the above equation), and how much more likely it is to get a true positive test result versus getting a false positive result (the ratio on the right hand side of the above equation).

Four steps are required to implement our method:

1. Treat each suggestive sign as a test of abuse, and combine the (as yet unknown) probabilities of abuse and of the occurrence of each suggestive sign using Bayes' rule (see Box 1). The assumptions being made when this model is created – in particular, whether the signs would be unrelated to each other in the absence of abuse – should be specifically identified, for later testing.
2. With the model as a guide, review the literature using appropriate definitions of the signs. This will determine if the suspect signs have previously been studied together and, if not, whether sufficient data exists on each individual sign to derive a probability of abuse using the model from 1 above. As probabilities are being estimated, studies must include an appropriate reference population. This criterion clearly excludes case studies and simple case series. Studies including rates of the suspect signs in the absence of abuse, as well as in their presence are also needed.
3. Extract the data needed to estimate each component probability for the model.
4. Run the model to obtain the conditional probability of abuse, given the suspicious signs being considered. With appropriate software, the uncertainty in each component sign's estimate will also be combined, so a credible interval may be calculated around the point probability estimate. The effect of different assumptions on the model should be tested using sensitivity analyses.

Full details of the development and application of this method are available elsewhere (Best et al. 2013). We concentrate here on the model's application to UK case

study introduced above, its findings, and their implications for the expert assessment of child abuse.

Step 1

The basic conjecture, that oronasal bleeding and an ALTE together predict abuse, can be expressed, using Bayes rule, as

$$\text{odds}(\text{abuse}|\text{bleed}, \text{ALTE}) = \text{odds}(\text{abuse}|\text{ALTE}) \times \frac{\text{Pr}(\text{bleed}|\text{abuse}, \text{ALTE})}{\text{Pr}(\text{bleed}|\text{no abuse}, \text{ALTE})}$$

This can be readily extended to include sibling death from SIDS as an additional sign. However, any association between sibling death and oronasal bleeding with ALTE might be due to abuse alone (conditional independence), or, additionally, to some other reason e.g., a hereditary disorder. The effects of these alternatives need to be modelled.

Step 2

ALTE was defined using current international guidelines (Michigan and King 1987). As no such guidance existed for oronasal bleeding, a set of relevant search terms was developed for the literature search. Details of the search strategy itself are available in (Best et al. 2013).

Step 3

Data about individual signs were extracted for the model's component probabilities, as set out at Table 1 below

Table 1: Probability numerators (n) and denominators (d) extracted from the available literature

<i>Study</i>	<i>d</i>	<i>n</i>	<i>Comments on data interpretation</i>
Pr(abuse ALTE)			
Truman & Ayoub 2002 (Truman and Ayoub 2002)	138	11	<ul style="list-style-type: none"> • Reports information on alive and dead children; • 6 confirmed abuse in dead children (although another 7 possible cases of suspicious death) • 5 reported abuse cases in live children (there is a possibility of 2 further cases; it is unclear if they are a subset of the other cases mentioned)
Pitetti et al 2002 (Pitetti et al. 2002)	128	3	<ul style="list-style-type: none"> • 1 further suspected case of abuse
(Altman et al. 2003)	243	6	<ul style="list-style-type: none"> • 6 cases of confirmed abuse • Another 4 possible cases of abuse, plus 35 further cases that are unlikely but can't be ruled out.
Altman et al 2003 (Gupta and Davies 2002)	65	2	<ul style="list-style-type: none"> • 2 confirmed cases of abuse • 15 cases with unknown diagnosis that may be abuse, but unlikely
Pr(bleed abuse, ALTE)			
Truman & Ayoub 2002 (Truman and Ayoub 2002)	6	1	<ul style="list-style-type: none"> • Confirmed abuse in children who died • 2 cases of suspected abuse have not been included as either abuse or not abused, but had experienced bleeding.
Truman & Ayoub 2002 (Truman and Ayoub 2002)	5	0 – 4	<ul style="list-style-type: none"> • It is not possible to determine exact numbers of abused children with bleeding in the live group, but the number must be between 0 and 4 of the cases initially identified. • If, when follow-up is included, the number of abused cases in the living group is as high as 7 then the number who had a bleed is between 0 and 6.
Southall et al 1997 (Southall et al. 1997)	37	10	<ul style="list-style-type: none"> • Cases #36 and #37 did not appear to have been abused, so were excluded from denominator.
Pr(bleed no abuse, ALTE)			
Truman & Ayoub 2002 (Truman and Ayoub 2002)	29	0	<ul style="list-style-type: none"> • Denominator was children who had died, with abuse neither confirmed nor suspected
Truman & Ayoub 2002 (Truman and Ayoub 2002)	98	5 – 9	<ul style="list-style-type: none"> • Children who had not died, excepting those reporting abuse. • Not possible to determine exact number of bleed cases amongst non-abused in live group but the total in the abused and non-abused groups in live children must be 5 through 9.

<i>Study</i>	<i>d</i>	<i>n</i>	<i>Comments on data interpretation</i>
Southall et al 1997 (Southall et al. 1997)	48	1	● Unclear whether or not cases #36 and #37 should be included as non-abused, so analyses were conducted both with and without these two cases in the denominator.

It can be seen that, sometimes, only a range of values could be deduced from a publication, and so the impact of this additional source of uncertainty should be modelled by sensitivity analyses. Only two papers (Truman and Ayoub 2002; Southall et al. 1997) provided information on sibling death through SIDS, but the problems of data display in (Truman and Ayoub 2002) did not allow even a range of values to be extracted. Of the 37 children with documented evidence of abuse in (Southall et al. 1997), 26 had siblings, of whom there were seven cases with one or more sibling deaths from SIDS. Two of these cases had also experienced bleeding from the nose or mouth, giving a direct estimate of 2/26 (8%) for the probability of bleeding and sibling death from SIDS in abused infants presenting with ALTE. Among the 48 cases who had not been abused, 45 had siblings, of whom there were two cases with a sibling death from SIDS. None of these cases also experienced oronasal bleeding, giving a direct estimate of the joint probability of oronasal haemorrhage and a previous sibling death from SIDS in non-abused children presenting with ALTE as 0/45 (0%).

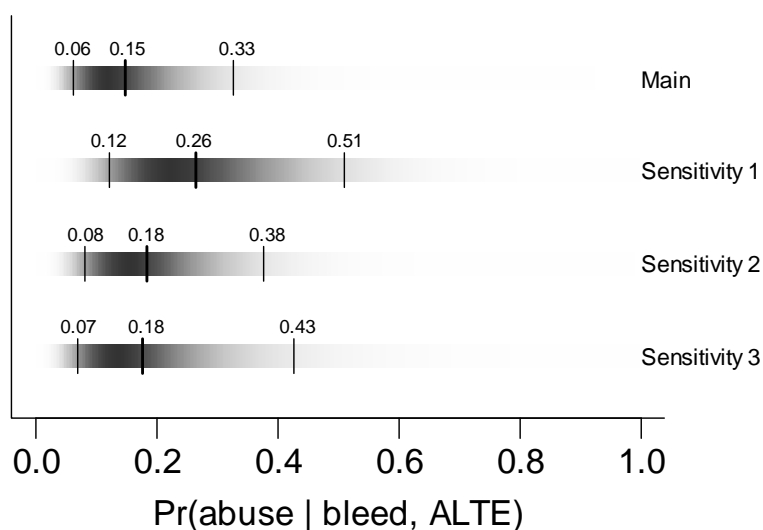
Step 4

Probability distributions for each component probability were estimated using Bayesian random effects meta-analysis of the data from Table 1, implemented in the software WinBUGS 1.4.3. These estimates were then combined using Bayes rule (see above) to obtain a probability distribution (from which credible intervals can be derived) for the probability of abuse. Full details of this are given in (Best et al. 2013).

Results

The most conservative (main) and three alternative analyses (sensitivity 1, 2 and 3) are presented in Figure 1: these address the uncertain data identified at Step 3. The main analysis took the lower number of bleed cases in abused infants with the upper number of bleed cases in non-abused infants in the study of (Truman and Ayoub 2002), and included possible or doubtful cases in the denominator. Sensitivity 1 took, instead, the upper number of bleed cases in abused infants and the lower number of bleed cases in non-abused infants. Sensitivity 2 counted possible abuse cases from (Truman and Ayoub 2002), (Pitetti et al. 2002) and (Altman et al. 2003). Sensitivity 3 additionally removed cases #36 and #37 in (Southall et al. 1997) from the denominator.

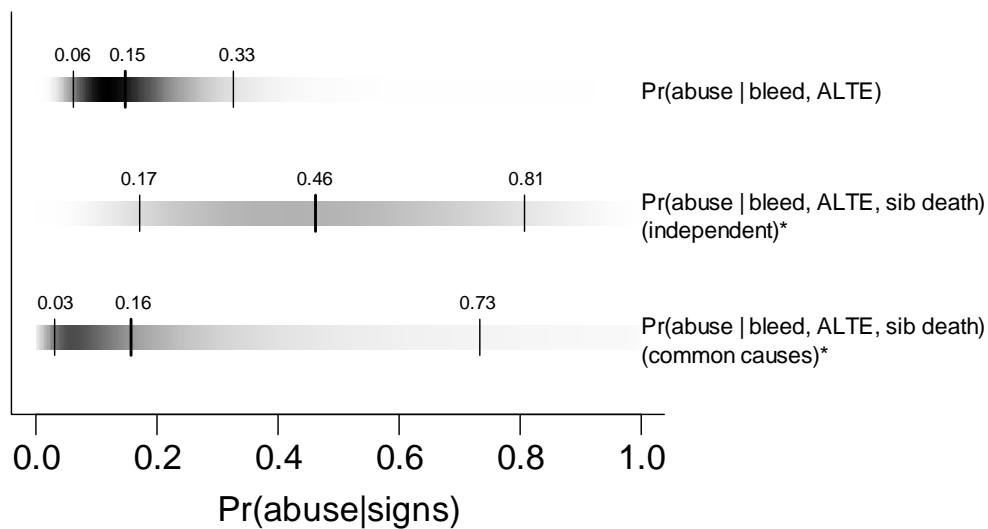
Figure 1: Probability of abuse in an infant presenting with ALTE and oronasal haemorrhage. Shaded strip indicates the plausible range of values for the probability of abuse taking into account the uncertainty in the available evidence; the darker the shading is, the more plausible the value. Outer vertical lines mark the 90% credible interval containing 90% of the plausible values; central vertical line marks the median of the plausible values and can be interpreted as the best estimate of the probability of abuse under each analysis.



When all the uncertainty in the data is considered, the range of reasonable probabilities for a combination of ALTE and oronasal bleeding predicting abuse lies between 0.06 and 0.51.

Figure 2 reports the impact of including sibling death from SIDS as an additional sign, as discussed in Step 1 above (results shown for main analysis only).

Figure 2: Probability of abuse in an infant presenting with ALTE, oronasal haemorrhage and sibling death from SIDS under different assumptions of independence. Shaded strip indicates the plausible range of values for the probability of abuse taking into account the uncertainty in the available evidence; the darker the shading is, the more plausible the value. Outer vertical lines mark the 90% credible interval containing 90% of the plausible values; central vertical line marks the median of the plausible values and can be interpreted as the best estimate of the probability of abuse under each analysis.



* independent = abuse is assumed to be only reason that bleed and sib death may be associated

* common causes = multiple common causes of bleed and sib death are assumed

The presence of sibling death from SIDS increased the probability of abuse, if abuse is the sole reason the two might be associated, from 0.15 (90% credible interval 0.06-0.33) to 0.46 (0.17-0.81). However, if multiple common causes are allowed, as is

likely in practice, then the median estimate of the probability of abuse does not increase, but there is far greater uncertainty about the plausible range of values (90% credible interval 0.03 to 0.73).

Discussion

Applying the method to the experts' conjecture in our case study generated two clinically important results.

1. In contrast to the expert view,, this combination of signs (oronasal bleeding and ALTE) can be no more than suggestive of abuse without further information, given the upper credible probability bound of 0.51, and its wide range. This difference is consistent with the impact of framing effects on expert opinion, as discussed above.
2. The extended model identified unhelpful additional information. Knowing about sibling death from SIDS simply increased the level of uncertainty when all reasonable assumptions about causation were considered, but this information risked generating a further framing effect, as expert practitioners are aware of the independent association of sibling death from SIDS with child abuse.

The first outcome shows that this method is able to construct an evidence-based estimate of the probability of abuse using sparse information, which may helpfully support expert judgments and guard against framing effects. The second, unexpected outcome suggests that it might also be valuable in distinguishing information, in this case sibling death, that is likely to confuse practitioners. While the example

demonstrated the potential to avoid an additional framing effect, this could also reduce the risk of false negatives through information overload or false assumptions.

The credible interval the model provides can also give experts guidance as to the strength of evidence from their own examinations needed to override the model's probability assessment. Thus it is consistent with the Structured Professional Judgment approach of (Bortoli and Dolan 2014), although differing by its estimation of a credible interval for the probability of abuse derived from the evidence. In the absence of such quantification, the use of guidelines seems not to improve the reliability of risk assessment when standardised information is provided (Skivenes and Stenberg 2013). As few as three risk factors has been shown to be effective in structuring child abuse reporting (Diderich et al. 2014) so, despite the much-discussed complexity of abuse, relatively simple models may prove sufficient to provide useful credible intervals, while the value of additional data may be assessed by addition to a simpler model, to see if it is likely to help or impair professional judgment.

Legal strengths of our method are that, as required in UK law (Thomas LJ, Beatson J, and Kitchin J 2010), it does not replace expert judgment with a statistically calculated opinion, and it also fits the "Daubert" recommendations for evidential support of expert opinion required in US law (Johnston and Sartwelle 2013).

This method does not infer causal connections from sign to abuse, though it may be adapted to do so under certain conditions (Dawid, Musio, and Fienberg 2013). Despite the risks of inadequate data, such meta-analyses may be the only way to accrue data in the absence of a dedicated study on the topic, and the sensitivity analyses undertaken (Best et al. 2013) suggest that, here, these risks were not sufficient to substantially bias the results.

In its prototype form, the method requires time, skills and resources not usually available to the practitioners who assess abuse. Development and evaluation of both simplified software and practitioner training would be needed before it could be generally recommended. Currently, its main application would seem to be in complex cases of suspected abuse, particularly when expert opinion is divided. In those cases, it could provide a much-needed “gold standard” for the evidential basis to support expert assessment. It could also be used as an academic tool, to identify where research is needed to clarify potential risk indicators, particularly in relation to base-rate information. The method could thus contribute significantly to the raising of both confidence and standards in expert diagnosis of child abuse worldwide.

References

- Altman, R. L., D. A. Brand, S. Forman, M. L. Kutscher, D. B. Lowenthal, K. A. Franke, and V. V. Mercado. 2003. "Abusive Head Injury as a Cause of Apparent Life-Threatening Events in Infancy." *Archives of Pediatrics and Adolescent Medicine* 157 (10): 1011.
- Best, N., D. Ashby, F. Dunstan, D. M. Foreman, and N. McIntosh. 2013. "A Bayesian Approach to Complex Clinical Diagnoses: A Case Study in Child Abuse." *Journal of the Royal Statistical Society Series A* 176 (1): 53–96.
- Bortoli, Lillian De, and Mairead Dolan. 2014. "Decision Making in Social Work with Families and Children: Developing Decision-Aids Compatible with Cognition." *British Journal of Social Work*, August, bcu087. doi:10.1093/bjsw/bcu087.
- Council for the Regulation of Healthcare Professionals v General Medical Council & Anor. 2005, 579 EWHC. Admin.
- Cradock, Gerald. 2011. "Thinking Goudge: Fatal Child Abuse and the Problem of Uncertainty." *Current Sociology* 59 (3): 362–78.
- Dawid, Philip, Monica Musio, and Stephen E. Fienberg. 2013. "From Statistical Evidence to Evidence of Causality." *arXiv Preprint arXiv:1311.7513*.
<http://arxiv.org/abs/1311.7513>.
- Diderich, Hester M., Mark Dechesne, Minne Fekkes, Paul H. Verkerk, Fieke D. Pannebakker, Mariska Klein Velderman, Peggy J. G. Sorensen, Simone E. Buitendijk, and Anne Marie Oudesluys-Murphy. 2014. "Facilitators and Barriers to the Successful Implementation of a Protocol to Detect Child Abuse Based on Parental Characteristics." *Child Abuse & Neglect* 38 (11): 1822–31.
doi:10.1016/j.chiabu.2014.07.016.
- Flaherty, Emalee G., Robert D. Sege, and Tammy Piazza Hurley. 2008. "Translating Child Abuse Research into Action." *Pediatrics* 122 (Supplement 1): S1–5.
- General Medical Council. 2008. "Determination on Dr David Southall"
http://paca.org.uk/pdf_files/notification_letter.pdf: 21st June 2011.

- González-Izquierdo, Arturo, Jenny Woodman, Lynn Copley, Jan van der Meulen, Marian Brandon, Deborah Hodes, Fiona Lecky, and Ruth Gilbert. 2010. "Variation in Recording of Child Maltreatment in Administrative Records of Hospital Admissions for Injury in England, 1997–2009." *Archives of Disease in Childhood* 95 (11): 918–25.
- Gupta, R., and F. Davies. 2002. "Apparent Life Threatening Events in Infants Presenting to an Emergency Department." *Emergency Medicine Journal* 19: 11–16.
- Ireland, Jane L. 2012. "Evaluating Expert Witness Psychological Reports: Exploring Quality." *Unpublished Report Presented to the Family Justice Council. February. Available at: [Http://www.Uclan.Ac.uk/news/FIANLVERSIONFEB2012.Pdf](http://www.Uclan.Ac.uk/news/FIANLVERSIONFEB2012.Pdf) (last Accessed on 11 January 2013).*
http://www.mfjc.co.uk/home/mfjccou1/public_ftp/resources/FINALVERSIONFEB2012.pdf.
- Johnson, S. R., G. A. Tomlinson, G. A. Hawker, J. T. Granton, and B. M. Feldman. 2010. "Methods to Elicit Beliefs for Bayesian Priors: A Systematic Review." *Journal of Clinical Epidemiology* 63 (4): 355–69.
- Johnston, James C., and Thomas P. Sartwelle. 2013. "The Expert Witness in Medical Malpractice Litigation Through the Looking Glass." *Journal of Child Neurology* 28 (4): 484–501.
- Kynn, M. 2008. "The 'heuristics and Biases' Bias in Expert Elicitation." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 171 (1): 239–64.
- Lichtenstein, S., P. Slovic, B. Fischhoff, M. Layman, and B. Combs. 1978. "Judged Frequency of Lethal Events." *Journal of Experimental Psychology: Human Learning and Memory* 4 (6): 551–78.
- Lindberg, Daniel M., Emily A. Blood, Kristine A. Campbell, Antoinette L. Laskey, and Rachel P. Berger. 2013. "Predictors of Screening and Injury in Contacts of Physically Abused Children." *The Journal of Pediatrics*.

- Mathews, Ben, Heather Payne, Catherine Bonnet, and David Chadwick. 2009. "A Way to Restore British Paediatricians' Engagement with Child Protection." *Archives of Disease in Childhood* 94 (5): 329–32.
- Meadow v General Medical Council. 2006, 146 EWHC. Admin.
- Michigan, D., and P. J. D. King. 1987. "National Institutes of Health Consensus Development Conference on Infantile Apnea and Home Monitoring, Sept 29 to Oct 1, 1986: CONSENSUS STATEMENT." *Pediatrics* 79 (2): 292–99.
- Ministry of Justice, and Department of Education. 2014. "Family Justice Reforms to Benefit Children - Press Releases - GOV.UK." April 22.
<https://www.gov.uk/government/news/family-justice-reforms-to-benefit-children>.
- Munro, E. 2010. "Learning to Reduce Risk in Child Protection." *British Journal of Social Work* 40 (4): 1135–51.
- Pitetti, RD, F. Maffei, K. Chang, R. Hickey, R. Berger, and MC Peirce. 2002. "Prevalence of Retinal Hemorrhages and Child Abuse in Children Who Present with an Apparent Life-Threatening Event." *Pediatrics* 110: 557–62.
- Sege, Robert D., and Emalee G. Flaherty. 2008. "Forty Years Later: Inconsistencies in Reporting of Child Abuse." *Archives of Disease in Childhood* 93 (10): 822–24.
- Seymour, Fred, Suzanne Blackwell, Sarah Calvert, and Briar McLean. 2014. "Counterintuitive Expert Psychological Evidence in Child Sexual Abuse Trials in New Zealand." *Psychiatry, Psychology and Law* 21 (4): 511–22.
- Skivenes, Marit, and Hanne Stenberg. 2013. "Risk Assessment and Domestic Violence – How Do Child Welfare Workers in Three Countries Assess and Substantiate the Risk Level of a 5-Year-Old Girl?" *Child & Family Social Work*, n/a – n/a.
doi:10.1111/cfs.12092.

- Southall, D. P., M. C. Plunkett, M. W. Banks, A. F. Falkov, and M. P. Samuels. 1997. "Covert Video Recordings of Life-Threatening Child Abuse: Lessons for Child Protection [see Comments]." *Pediatrics* 100 (5): 735–60.
- Thomas LJ, Beatson J, and Kitchin J. 2010. "R v T." *EWCA Crim* 2439.
- Truman, TL, and CC Ayoub. 2002. "Considering Suffocatory Abuse and Munchausen by Proxy in the Evaluation of Children Experiencing Apparent Life-Threatening Events and Sudden Infant Death Syndrome." *Child Maltreatment* 7: 138–48.
- Tversky, A., and D. Kahneman. 1981. "The Framing of Decisions and the Psychology of Choice." *Science* 211 (4481): 453–58.
- White, Sue, David Wastell, Karen Broadhurst, and Chris Hall. 2010. "When Policy O'erleaps Itself: The 'tragic Tale' of the Integrated Children's System." *Critical Social Policy* 30 (3): 405–29. doi:10.1177/0261018310367675.