

22 January 2001
Original version

A case-control study of autism and mumps-measles-rubella vaccination using the general practice research database: design and methodology

Liam Smeeth [liam.smeeth@lshtm.ac.uk]
Andrew J Hall [andy.hall@lshtm.ac.uk]
Eric Fombonne [e.fombonne@iop.kcl.ac.uk]
Laura C Rodrigues [laura.rodrigues@lshtm.ac.uk]
Xiangning Huang [xiangning.huang@lshtm.ac.uk]
Peter G Smith [peter.smith@lshtm.ac.uk]

A case-control study of autism and mumps-measles-rubella vaccination using the general practice research database: design and methodology

Liam Smeeth¹ (liam.smeeth@lshtm.ac.uk)
Andrew J Hall² (andy.hall@lshtm.ac.uk)
Eric Fombonne³ (e.fombonne@iop.kcl.ac.uk)
Laura C Rodrigues² (laura.rodrigues@lshtm.ac.uk)
Xiangning Huang² (xiangning.huang@lshtm.ac.uk)
Peter G Smith² (peter.smith@lshtm.ac.uk)

¹ Department of Epidemiology and Population Health

² Department of Infectious and Tropical Diseases

London School of Hygiene and Tropical Medicine

Keppel Street, London WC1E 7HT, UK

³ Institute of Psychiatry/King's College London
Department of Child and Adolescent Psychiatry
MRC Child Psychiatry Unit
Denmark Hill
16 De Crespigny Park
London SE5 8AF, UK

Correspondence to: Dr Liam Smeeth

Tel: ++44 (0)20 7927 2296

Fax: ++44 (0)20 7580 6897

A case-control study of autism and mumps-measles-rubella vaccination using the general practice research database: design and methodology

Abstract

Background

An association between mumps-measles-rubella (MMR) vaccination and the onset of symptoms typical of autism has recently been suggested. This has led to considerable concern about the safety of the vaccine.

Methods

A matched case-control study using data derived from the United Kingdom General Practice Research Database. Children with a possible diagnosis of autism will be identified from their electronic health records. All diagnoses will be validated by a detailed review of hospital letters and by using information derived from a parental questionnaire. Ten controls per case will be selected from the database. Conditional logistic regression will be used to assess the association between MMR vaccination and autism. In addition case series analyses will be undertaken to estimate the relative incidence of onset of autism in defined time intervals after vaccination. The study is funded by the United Kingdom Medical Research Council.

Discussion

Electronic health databases offer tremendous opportunities for evaluating the adverse effects of vaccines. However there is much scope for bias and confounding. The rigorous validation of all diagnoses and the collection of additional information by parental questionnaire in this study are essential to minimise the possibility of misleading results.

Background

The epidemiology of autism

Autism is a pervasive developmental disorder characterised by abnormalities in the development of language, communication abilities, and social interactions and by a pattern of restricted play and behaviour which tends to be highly repetitive, unimaginative and rigid.[1] By definition, the abnormalities must be present by the age of three years, although the diagnosis is usually not made until the age of four or five years.[2] In studies of the consistency of diagnosis there has been a high consensus between psychiatrists and coding instruments.[3]

The age at which parents first recognise an abnormality is variable, with 40% of autistic children having shown typical features by the age of one year and most by the age of two years.[4] This age is influenced by the degree of associated mental retardation and birth order (the less severe and first born children tending to have later age of parental recognition).[5] Most population-based studies have found a prevalence of autism between 5 and 10 per 10,000 children.[6]

MMR vaccination and autism

In 1998 a link was suggested between mumps-measles-rubella (MMR) vaccination and autism.[7] This was based on an uncontrolled case series of 12 children referred to a paediatric gastroenterology unit with a history of normal development followed by loss of acquired skills, including language, together with diarrhoea and abdominal pain. It was suggested that the gastrointestinal and developmental symptoms were a syndrome that could have been triggered by MMR vaccination. The study was widely criticised[8,9] but generated considerable media interest[10] and led to a small fall in MMR coverage in the United Kingdom.[11] A larger case series of 60 children with the same combination of clinical findings has recently been published.[12]

Since the first study by Wakefield et al, a number of published studies have looked specifically at this issue. In a small study from Finland, among 31 children who had reported a gastrointestinal adverse reaction to MMR vaccination, none had subsequently developed signs of autism.[13] A similar larger study looked at all notified serious adverse events following MMR vaccination in Finland over a 14 year

period.[14] There were no new cases of autism among 173 notified adverse events. However such routine passive surveillance systems have a number of weaknesses for epidemiological studies.[15] There is no control group, the quality of the data may be sub-optimal and detecting an effect depends entirely on clinicians believing a new illness was due to vaccination. In Sweden no increase was apparent in the incidence of autism following the introduction of MMR vaccination.[16] Both these studies included small numbers of children with autism and had limited ability to assess the link between MMR vaccine and autism. The United Kingdom Committee on Safety of Medicines set up a working party to assess parental and medical reports of children who had developed autism, Crohn's disease or similar disorders following MMR vaccination. The Working Party Report was, by its own description, solely a descriptive account of those children whose parents had sought legal advice about possible vaccine damage.[17] The Report highlighted bias in the way affected children were selected for inclusion in the study and the lack of any control group before concluding that they could not prove or refute the suggested associations between MMR vaccine and autism. A single large high quality epidemiological study has been published.[18] This study included 293 children with confirmed autism from North Thames health districts. From time series trends analysis, age of diagnosis in vaccinated and unvaccinated groups and a case series analysis, the authors concluded there was no evidence to support an association. The study did find a positive association between MMR vaccination and first parental concerns in the first six months following vaccination. Although the authors considered that this finding was likely to be either a chance finding or due to inaccuracy in recalling the date of onset of symptoms, this interpretation has been disputed.[19] It was also suggested that because the study only considered a date of onset up to six months after vaccination, a causal link may have been missed.[20] The authors of the study have undertaken a re-analysis looking at longer post vaccination risk periods, and again found no evidence to support a link between MMR vaccination and autism.[21]

In the light of continuing concern about the proposed link between MMR vaccination and autism[22-26] we plan to undertake a case-control study using data derived from the General Practice Research Database.

Objectives

The study has two linked objectives with respect to MMR vaccination. Firstly to determine if autistic children are more likely to have received MMR vaccine prior to disease onset. Secondly to examine whether there is any association between clinical onset of disease and the timing of the MMR vaccination.

Materials and methods

The General Practice Research Database (GPRD)

The GPRD (previously known as the VAMP Research Bank) was set up in 1987 and is now held by the Medicines Control Agency.[27,28] It contains complete prescribing and diagnostic information from a large number of general practices and is the largest source of continuous data on illness and prescribing habits in the United Kingdom. Over 200 published studies have been completed using the database. Participating general practitioners were given instruction over a 12-month period regarding standardised recording of clinical information into their computing systems. The general practices are broadly representative of all practices in the United Kingdom in terms of geographical distribution and size and the age and sex distributions of the population included in the GPRD are similar to the whole United Kingdom population.[29] The data available directly from the database include all drug prescriptions and their indication, a record of every consultation and of every diagnosis. The data collected is audited regularly and the participating general practices are subjected to a number of quality checks. Of the practices contributing to the database, about 280 practices, with a combined population of around 2.1 million patients currently pass these rigorous quality checks. The quality of the information in the database has been validated in a number of independent studies and has been found to be high.[30-35]

The general practitioners keep all referral letters, hospital discharge summaries and other clinically relevant letters in a manual file. In addition to the electronic health record, questionnaires can be sent to patients (or their parents) via general practitioners, and copies letters relating to referrals and hospital care can be obtained. The data are held anonymously in the central GPRD database, with patient identifiers removed.

Identification of affected children

Children with putative autism will be identified by searching the whole electronic record of all people included in the GPRD for diagnostic codes which possibly relate to a diagnosis of autism. MMR vaccine was introduced in the United Kingdom for all children aged 12 to 15 months in October 1988. An MMR catch-up campaign was also launched for older children in 1988. We will separately identify those children with putative autism born after and before mid-1987, which separates out those children likely to have received the MMR vaccine around the age of 1 year and those likely to have received it at a later date. Separate analyses will be conducted on these two groups. Although all major past diagnoses are recorded in practice computers when new patients register with practices, such recording may be incomplete. To overcome this potential problem, we will identify children first diagnosed when they were registered with practices participating in the GPRD. Children diagnosed prior to registration with the GPRD will be analysed separately with their matched controls. The results from these two groups will be pooled if they are similar.

Identification and selection of controls

For each affected child we will sample matched controls from the GPRD. Two groups of controls will be selected. The first group will consist of five people with no record of autism matched on age (\pm one year), sex and practice. Matching in this group aims to control for possible confounding by the general practice participants are registered with. The second group will be of similar size and will be matched on age and sex but not on practice, to avoid the possibility of overmatching. For children diagnosed while registered with a GPRD practice, the date of diagnosis will be called the index date. The controls will be selected from those patients registered with the GPRD on the index date of the affected child they are matched to. We will not be able to apply the same method for selecting controls for children with autism diagnosed prior to registering with a practice participating in the GPRD because they will not have an index date. Therefore the matched controls for children diagnosed prior to registering with a practice participating in the GPRD will be selected from all patients registered with the GPRD on the date the affected child registered with the GPRD.

Questionnaire to parents of affected children and controls

A questionnaire will be sent to the parents of all affected children and to two controls per affected child, one matched on practice and one not matched on practice and closest in age to the affected child. The questionnaire to parents of children with autism will include an autism screening questionnaire[36] and will solicit information on: the date of first symptoms of autism and earliest date of parental concern about symptoms possibly related to autism; the educational status of the child; the knowledge and beliefs of parents regarding the causes of autism; and family history of pervasive developmental problems. In addition the questionnaire will specifically ask about family history of pervasive developmental problems, genetic disorders and about regression (loss of skills) allowing us to classify affected children into those with reported regression and those without.

For both affected children and controls the questionnaire will include questions about: the socioeconomic status of the parents; birth order and family size; history of bowel disturbance in the child; and vaccination history.

Diagnosis: definition and validation

As a first step to validate the diagnoses, copies of all hospital summaries will be requested from the GPs concerned. Previous studies using the GPRD have obtained full copies of hospital summaries on over 90% of patients still registered with a collaborating practices.[35,37] We will obtain copies of letters relating to both autism and to all other reasons for hospital investigation or attendance, including bowel investigations and inflammatory bowel disease (see below). The basis for the diagnosis of autism, evidence of associated genetic disorders and the date of first attendance for possible autism will be extracted from the records.

There is strong agreement among child psychiatrists about concepts of and operational definitions for autism.[3] We believe that no child will be labelled as autistic in the GP record without referral to child psychiatry services. Two studies have specifically documented the completeness of the information in the GPRD about referrals occurring and their outcome.[30,31]

All information about children possibly affected by an autistic spectrum disorder, including information about the current educational status of the child from the questionnaire, will be reviewed independently by two child psychiatrists. They will use DSM-IV / ICD 10 research diagnostic criteria to define autistic

spectrum disorders, and will attempt to subtype the disorders according to their phenomenology. In particular they will separate and sub-classify autistic disorder in DSM-IV or childhood autism in ICD-10, Asperger's disorder, atypical autism / pervasive developmental disorder not otherwise specified, and other forms of pervasive developmental disorders (i.e. Rett's syndrome and childhood disintegrative disorder). This will be achieved by rating the developmental abnormalities on a symptom basis and then applying diagnostic algorithms. They will also make an overall global judgement about the clinical pattern and rate their confidence in this final diagnostic judgement in order to allow for difficult or improbable diagnoses to be treated separately. Inter-rater reliability estimates will result from this exercise. Separate analyses will be carried out for children with a definite diagnosis and for children with a definite or probable diagnosis in order to assess the potential impact of misclassification.

Exclusion of affected children with an alternative aetiology

Inclusion of affected children who have an established alternative aetiology may bias the estimated odds ratio for the association between vaccination and adverse outcome towards unity.[38] Some children will have medical disorders thought to have a causal association with autism (fragile X disorder, tuberous sclerosis, phenylketonuria, congenital rubella) and will be excluded. A recent review estimated that this will lead to the exclusion to at most 6% of affected children.[6]

Determination of date of onset

From the GP record, hospital letters and parental questionnaire for each affected child we will extract the date of:

- first attendance to the GP with symptoms or problems potentially relating to a future diagnosis of autism, such as behavioural difficulties (e.g. sleeping or eating difficulties), delay in motor development and milestones, delay in language development, abnormalities in social development (for example delayed smiling, lack of reciprocity, lack of anticipation, odd behaviours);
- first concerns or symptoms as recorded in the hospital letters;

- definitive diagnosis from the hospital letters;
- first parental concern of symptoms of autism collected retrospectively.

The first three dates will be based on existing records and both the date and the relationship of the date to the timing of MMR vaccination will not be affected by errors of memory. First parental concerns about autism may have occurred many years ago and some error in accurately remembering the exact date is to be expected. In addition, it is possible that parental recall of the date of onset of symptoms relative to the timing of MMR vaccination may be affected by the recent publicity about a possible link between MMR and autism. The proposed link between MMR vaccine and autism was first publicised in February 1998. After this date public and media concern about the possible link may have affected the likelihood of a child attending the GP with problems and in particular the timing of the presentation relative to MMR vaccination. Children with a date of first symptoms after February 1998 will be analysed separately to assess the effect of possible bias.

For the main analyses the date of onset will be the earliest of either the date of first attendance to the GP with symptoms potentially relating to a future diagnosis of autism or the date of first concerns or symptoms as recorded in the hospital letters.

Assessment of exposure

Exposure to MMR vaccine will be extracted from the GP electronic record. This method has a two advantages. Firstly it will avoid recall bias either about vaccine status or about the timing of vaccination relative to the onset of symptoms. Secondly there are good reasons to expect the vaccine data to be complete. All general practitioners participating in the GPRD undertake to include all medications prescribed or administered in the computerised record. In addition, United Kingdom general practitioners have a financial incentive to accurately record childhood vaccination status. Finally there is excellent agreement between prescribing data from the GPRD and national data from the Prescription Pricing Authority.[34]

Analyses

Case-control

Conditional logistic regression will be used to undertake matched case-control analyses. Potential confounding factors include presence of another chronic illness, social class, birth order, and family size. We will initially undertake a series of univariate analyses. Factors that appear to be associated with autism ($P < 0.2$) will be carried forward to a multivariate model. Likelihood ratio tests will be used for all tests of significance. Two analyses will be carried out. The first will estimate the odds ratio for the development of symptoms in specific time periods after vaccination with MMR. This method provides an alternative approach to the case series approach outlined below. The second will assess exposure to MMR vaccine at any time prior to symptom development. This analysis differs from the case series approach in that no assumption is required about the likely interval from vaccination to disease onset if there is a causal association.

Case series

The case series uses data on affected children only to estimate the relative incidence of clinical events in a defined interval after vaccination compared to time periods outside this defined interval.[39,40] The method has been used to estimate the relative incidence of febrile convulsions following DTP and MMR vaccines[41] and was also used in a recent study of the onset of autism following MMR vaccine.[18] We will examine periods of 1 month, 2 months, 4 months, 6 months and 1 and 2 years after vaccination. The reference period for each individual will consist of every month from birth up until February 1998, which was when the possible link between MMR vaccine and autism became widely known, excluding the post-vaccination period being studied. All analyses will be finely stratified for age, the exact stratification will depend on the age distribution of the affected children.

The two approaches estimate different parameters. The case control approach will estimate the odds ratio for whether children who are vaccinated have an increased chance overall of developing autism than children who are not vaccinated. The case series will estimate the relative incidence of autism in the period following MMR vaccination.

Power

We estimate we will be able to include a minimum of 400 children with a diagnosis of autism in the analyses. Over the entire study period we estimate the proportion of children in the control group who will have received MMR vaccination to be around 85%.[11] With 5 controls per affected child in the case-control analysis we will be able to detect the following minimum odds ratios for the association between autism and MMR vaccination with 90% power at the 5% significance level: 1.8 if average MMR coverage among controls is 85%, or 2.0 if average MMR coverage among controls is 90%. For the case series analysis assuming an 85% vaccine coverage rate (a conservative estimate), we will have 90% power at the 5% significance level to detect a minimum relative incidence for autism of 1.6 in the 1 month following MMR vaccine.

Ethical approval

The Scientific and Ethical Advisory Group is a central ethical committee specially set up by the Department of Health to oversee use of the GPRD. They have approved the study as have the ethics committee of the London School of Hygiene and Tropical Medicine. The use of confidential patient data in this study is fully within the recent guidelines from both the United Kingdom Medical Research Council[42] and the General Medical Council[43] about the use of personal information in medical research.

Discussion

Electronic databases offer several important advantages for epidemiological studies of adverse events from vaccination. All people affected by the adverse event (or a random sample) can be drawn from existing records, usually avoiding the problem of ascertainment being linked to exposure, although bias may not entirely be removed if people affected were diagnosed after the hypothesis was known. As controls can be sampled from all other participants in the database, biased selection of controls is less likely to occur.

Records of date of vaccination and onset of symptoms, are also less likely to be biased, in particular if they precede the hypothesis coming into public domain.

The major disadvantage of such databases is that data quality and completeness may not always be optimal. In particular, all diagnoses of autism will not have been made using the same criteria applied in a consistent manner.

Vaccines are without doubt among the most effective public health interventions, but thorough investigation of suspected adverse effects is necessary. Case-control studies using electronic health databases offer a uniquely efficient method for evaluating adverse effects of vaccines. However they also offer scope for bias and confounding to produce misleading results. The rigorous validation of all possible diagnoses and the collection of additional information by parental questionnaire in this study will be both time consuming and expensive, but we view this as essential to minimise the possibility of biased results.

Funding

The study is funded by the United Kingdom Medical Research Council.

Liam Smeeth is funded by a Research Fellowship from Thames NHS Research and Development Directorate.

Reference List

1. World Health Organisation: *The ICD-10 classification of mental and behavioural disorders*. Geneva: World Health Organisation; 1992.
2. Howlin P, Moore A: **Diagnosis in autism**. *Autism* 1997, **1**: 135-162.
3. Volkmar FR, Klin A, Siegel B, Szatmari P, Lord C, Campbell M *et al.*: **Field trial for autistic disorder in DSM-IV**. *Am J Psychiatry* 1994, **151**: 1361-1367.
4. Rogers SJ, DiLalla DL: **Age of symptom onset in young children with pervasive developmental disorders**. *J Am Acad Child Adolesc Psychiatry* 1990, **29**: 863-872.

5. De Giacomo A, Fombonne E: **Parental recognition of developmental abnormalities in autism.** *Eur Child Adolesc Psychiatry* 1998, **7**: 131-136.
6. Fombonne E: **The epidemiology of autism: a review.** *Psychol Med* 1999, **29**: 769-786.
7. Wakefield AJ, Murch SH, Anthony A, Linnell J, Casson DM, Malik M *et al.*: **Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children.** *Lancet* 1998, **351**: 637-641.
8. Nicoll A, Elliman D, Ross E: **MMR vaccination and autism.** *BMJ* 1998, **316**: 715-716.
9. Chen RT, DeStefano F: **Vaccine adverse events: causal or coincidental?** *Lancet* 1998, **351**: 611-612.
10. Begg N, Ramsay M, White J, Bozoky Z: **Media dents confidence in MMR vaccine.** *BMJ* 1998, **316**: 561.
11. Thomas DR, Salmon RL, King J: **Rates of first measles-mumps-rubella immunisation in Wales (UK).** *Lancet* 1998, **351**: 1927.
12. Wakefield AJ, Anthony A, Murch SH, Thomson M, Montgomery SM, Davies S *et al.*: **Enterocolitis in children with developmental disorders.** *Am J Gastroenterol* 2000, **95**: 2285-2295.
13. Peltola H, Patja A, Leinikki P, Valle M, Davidkin I, Paunio M: **No evidence for measles, mumps, and rubella vaccine-associated inflammatory bowel disease or autism in a 14-year prospective study.** *Lancet* 1998, **351**: 1327-1328.
14. Patja A, Davidkin I, Kurki T, Kallio MJ, Valle M, Peltola H: **Serious adverse events after measles-mumps-rubella vaccination during a fourteen-year prospective follow-up.** *Pediatr Infect Dis J* 2000, **19**: 1127-1134.
15. Ellenberg SS, Chen RT: **The complicated task of monitoring vaccine safety.** *Public Health Rep* 1997, **112**: 10-20.

16. Gillberg C, Hejbel H: **MMR and autism.** *Autism* 1998.
17. Committee on Safety of Medicines. Report of the Working Party on MMR Vaccine. 1999. London, Committee on Safety of Medicines.
18. Taylor B, Miller E, Farrington CP, Petropoulos MC, Favot-Mayaud I, Li J *et al.*: **Autism and measles, mumps, and rubella vaccine: no epidemiological evidence for a causal association.** *Lancet* 1999, **353**: 2026-2029.
19. Wakefield AJ: **MMR vaccination and autism [letter].** *Lancet* 1999, **354**: 949-950.
20. Roger JH: **The MMR question.** *Lancet* 2000, **356**: 160-161.
21. Farrington CP, Miller E, Taylor B: **MMR and autism: further evidence against a causal association.** *Vaccine* 2000.
22. Loff B, Cordner S: **Australia's measles campaign challenged.** *Lancet* 1998, **352**: 1368.
23. Tookey PA, Peckham CS: **Surveillance of congenital rubella in Great Britain, 1971-96.** *BMJ* 1999, **318**: 769-770.
24. Burton D. Opening statement by the Chairman. Government Reform Committee United States House of Representatives. Autism: Present Challenges, Future Needs - Why the Increased Rates? Committee on Government Reform. 6th April 2000. Washington DC.
(<http://www.house.gov/reform/hearings/healthcare/00.06.04/>)
25. Pareek M, Pattison HM: **The two-dose measles, mumps and rubella (MMR) immunisation schedule: factors affecting maternal intention to vaccinate.** *Br J Gen Pract* 2000, **50**: 969-971.
26. Petrovic M, Roberts R, Ramsay M: **Second dose of measles, mumps, and rubella vaccine: questionnaire survey of health professionals.** *BMJ* 2001, **322**: 82-85.

27. Walley T, Mantgani A: **The UK General Practice Research Database.** *Lancet* 1997, **350**: 1097-1099.
28. Lawson DH, Sherman V, Hollowell J: **The General Practice Research Database. Scientific and Ethical Advisory Group.** *Q J Med* 1998, **91**: 445-452.
29. Office for National Statistics: *Key health Statistics from general practice 1996 (Series MB6 No. 1).* London: Office for National Statistics; 1998.
30. Jick H, Jick SS, Derby LE: **Validation of information recorded on general practitioner based computerised data resource in the United Kingdom.** *BMJ* 1991, **302**: 766-768.
31. Jick H, Terris BZ, Derby LE, Jick SS: **Further validation of information recorded on general practitioner based computerised data resource in the United Kingdom.** *Pharmacoepid Drug Safety* 1992, 347-349.
32. Nazareth I, King M, Haines A, Rangel L, Myers S: **Accuracy of diagnosis of psychosis on a general practice computer system.** *BMJ* 1993, **307**: 32-34.
33. van Staa T, Abenhaim L: **The quality of information recorded on a UK database of primary care records: a study of hospitalisations due to hypoglycaemia and other conditions.** *Pharmacoepid Drug Safety* 1994, **3**: 15-21.
34. Hollowell J: **The General Practice Research Database: quality of morbidity data.** *Popul Trends* 1997, 36-40.
35. Lawrenson R, Todd JC, Leydon GM, Williams TJ, Farmer RD: **Validation of the diagnosis of venous thromboembolism in general practice database studies.** *Br J Clin Pharmacol* 2000, **49**: 591-596.
36. Berument SK, Rutter M, Lord C, Pickles A, Bailey A: **Autism screening questionnaire: diagnostic validity.** *Br J Psychiatry* 1999, **175**: 444-451.

37. Garcia Rodriguez LA, Ruigomez A, Jick H: **A review of epidemiologic research on drug-induced acute liver injury using the general practice research data base in the United Kingdom.** *Pharmacotherapy* 1997, **17**: 721-728.
38. Wentz KR, Marcuse EK: **Diphtheria-tetanus-pertussis vaccine and serious neurologic illness: an updated review of the epidemiologic evidence.** *Pediatrics* 1991, **87**: 287-297.
39. Farrington CP: **Relative incidence estimation from case series for vaccine safety evaluation.** *Biometrics* 1995, **51**: 228-235.
40. Farrington CP, Nash J, Miller E: **Case series analysis of adverse reactions to vaccines: a comparative evaluation.** *Am J Epidemiol* 1996, **143**: 1165-1173.
41. Farrington P, Pugh S, Colville A, Flower A, Nash J, Morgan-Capner P *et al.*: **A new method for active surveillance of adverse events from diphtheria/tetanus/pertussis and measles/mumps/rubella vaccines.** *Lancet* 1995, **345**: 567-569.
42. Medical Research Council: *Personal information in medical research*. London: MRC; 2000.
43. General Medical Council: *Confidentiality*. London: GMC; 2000.

24 January 2001
Reviewers' reports

A case-control study of autism and mumps-measles-rubella vaccination using the general practice research database: design and methodology

Liam Smeeth [liam.smeeth@lshtm.ac.uk]
Andrew J Hall [andy.hall@lshtm.ac.uk]
Eric Fombonne [e.fombonne@iop.kcl.ac.uk]
Laura C Rodrigues [laura.rodrigues@lshtm.ac.uk]
Xiangning Huang [xiangning.huang@lshtm.ac.uk]
Peter G Smith [peter.smith@lshtm.ac.uk]

Charles Hoff

Comments

1. This is a well-written paper outlining the use of a large database to address some questions about a possible association between 'exposure' to the MMR vaccine and risk for eventual development of autism — a topic of considerable interest and disagreement. There are, however, some minor syntactical issues which should be addressed by the authors or editorial staff. If some of these are due to differences in practice between U.S. and U.K. English, then ignore them.

- a) p. 9., line 4: there should be an 'of' between the words 'copies' and 'letters',
- b) p 10., line 5: dangling prep.; I suggest saying "practice with whom participants are registered.",
- c) p10, line 10: ditto; I suggest saying "child with whom they are matched.",
- d) p. 11, line 3: there should be a comma between 'addition' and 'the',
- e) p. 11, line 6; dangling prep.; I suggest saying "with and those without reported regression.",
- f) p13, line 1: "exclusion to at" should be "exclusion of at",
- g) p. 13, line 7: there should be a comma after 'e.g.',
- h) p. 14, line 5: ditto comma after 'analyses',
- i) p. 14, line 10: "has a two" should be "has two",
- j) p .14, lines 11, 12, 13: there should be commas after 'Firstly', 'Secondly', and 'Finally' — respectively,
- k) p. 17, line 12 and 20: ditto after 'particular' and 'However',
- l) throughout the ms. 'case series' should be 'case-series' and 'case control' should be 'case-control'.

m) there are possibly more places where commas are required which I missed.

2. The proposed use of conditional logistic regression seems appropriate. However, since there is the issue of age and elapsed time from vaccination to when the patient's status is examined, survival analytic approaches would also seem appropriate (e.g., stepwise Cox regression). I believe that this would add more 'robustness' in clarifying associations between vaccination and outcome. My logic here is that there is the possibility that some younger examined patients with short elapsed times might yet develop autism after the study is completed. I encourage the authors to discuss this with an epidemiologic-oriented biostatistician. I have not checked this in the cited papers but I will bet 'dollars to doughnuts' that none of the studies cited used this approach. This might give the authors 'a leg up' in their investigation!!!

I await their findings with great anticipation. I suggest accepting this ms. with the syntactical revisions (if appropriate). The possible use and discussion of survival analysis, I leave to the authors.

Competing interests: none declared

CP Farrington

General comments

This paper describes the methods to be used in a study of a possible causal association between MMR vaccination and autism using the UK GPRD database. In view of the controversy surrounding this issue, the paper is clearly topical, though somewhat unsatisfying as no data are presented: reading it was rather like staring at an enticing menu outside a closed restaurant. I look forward to seeing the results.

Investigating a possible causal association between MMR vaccination and autism is fraught with difficulty. The overall strategy proposed by the authors seems sound, though they do not address some key problems of bias and power with their approach.

Major comments

1. The case control study will estimate the odds ratio of autism in exposed and non-exposed subjects, where exposure is defined as MMR vaccination prior to the index date. This relies on vaccinated and unvaccinated subjects being comparable in their underlying risk of developing autism. When vaccination coverage is very high, as is the case for children born after 1987, this cannot be assumed to be true, as the unvaccinated group will be highly selected. This potentially serious bias is not mentioned by the authors.

2. It is not clear from the design of the study that individual-level confounding will be adequately dealt controlled for. A review of likely confounders would have been useful in order to assess whether matching on practice will be sufficient, which I personally doubt.

3. The first group of controls are matched on practice and sex but only roughly on age, within 1 year. Does this mean that a 1 year-old case could be matched to a 2

year-old control, or vice versa? If so this could bias the results as MMR vaccination is very age-dependent.

4. The authors propose to undertake the case-control study with 'ever vaccinated' as the risk factor. It might be as well to plan analyses in which exposure is taken as vaccination within fixed periods prior to the index date, as I would suspect that the power of the study might otherwise be very low. The design assumption of 85% vaccination coverage is unrealistically low in the cohort born after 1987: the coverage by 2 years is more likely to be closer to 95%.

Minor comments

1. p7 lines 11-13. In the Taylor et al study, risk periods of up to 2 years were investigated, not just 6 months.

2. p7 line 16. Delete journal name and date from reference 21: the paper has not yet been accepted for publication.

3. p14 lines 13-14. The authors state that 'there are good reasons to expect the vaccine data to be complete'. Will this be checked?

4. p14 line -1. Will separate analyses be undertaken for the two sets of controls? If so, how will the authors interpret discordant results?

5. p15 lines 9-11. The case series method does not necessarily require a defined post-vaccination risk period, as shown by reference 21, and sidesteps the problem of individual-level confounding.

Level of interest

A paper important to those with closely related research interests

Quality of written English

Acceptable

Competing interests

Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this paper?

No

Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this paper?

No

Do you have any other financial competing interests?

No

Are there any non-financial competing interests you would like to declare in relation to this paper?

No

Open peer review

Do you consent to making your signed report accessible on the website should the paper be published?

Yes

12 February 2001

Authors' response to referees' comments

A case-control study of autism and mumps-measles-rubella vaccination using the general practice research database: design and methodology

Liam Smeeth [liam.smeeth@lshtm.ac.uk]

Andrew J Hall [andy.hall@lshtm.ac.uk]

Eric Fombonne [e.fombonne@iop.kcl.ac.uk]

Laura C Rodrigues [laura.rodrigues@lshtm.ac.uk]

Xiangning Huang [xiangning.huang@lshtm.ac.uk]

Peter G Smith [peter.smith@lshtm.ac.uk]

The referees comments are in italics, our responses are given in plain text.

Review by C. Paddy Farrington

Major comments

1. The case control study will estimate the odds ratio of autism in exposed and non-exposed

subjects, where exposure is defined as MMR vaccination prior to the index date.

This relies on vaccinated and unvaccinated subjects being comparable in their underlying

risk of developing autism. When vaccination coverage is very high, as is the case for children born after 1987, this cannot be assumed to be true, as the unvaccinated group

will be highly selected. This potentially serious bias is not mentioned by the authors.

2. It is not clear from the design of the study that individual-level confounding will be adequately dealt controlled for. A review of likely confounders would have been useful in

order to assess whether matching on practice will be sufficient, which I personally doubt.

Both these comments centre on the issue of confounding. We agree this is a crucial issue in this (and similar) studies. For this reason we are supplementing information held in the GP record with a questionnaire to all affected children and to two controls per affected child. We have expanded the discussion of confounding in the paper along the lines suggested by the referee.

3. The first group of controls are matched on practice and sex but only roughly on age,

within 1 year. Does this mean that a 1 year-old case could be matched to a 2 year-old control, or vice versa? If so this could bias the results as MMR vaccination is very age-dependent.

Controls will be selected to be as close in age as possible to the cases. For the control group not matched on practice, controls will be selected from the whole population of the GPRD. The first control selected will be the child of the same sex with no diagnosis of autism whose date of birth is the nearest to the date of birth of the case. The second control will be the child with the next nearest date of birth. This will continue until we have selected five cases. The procedure will be repeated for the control group matched on practice, with children selected from the same practice. It is therefore likely that controls will be very close in age to the cases. The 1 year limit refers to the maximum difference in age we will accept: if less than 5 controls (in either group) are found within 1 year of age of a case, we will not select any further controls for that group and for that case.

In a recent study using the same methodology, we found the mean difference in age between cases and two age matched controls to be less than 4 days.

We will assess how well matched on age the cases and controls for each group. As now stated in the paper, we will also examine the effects of this matching in our analysis: comparing the results for those children very closely matched on age (for example within 6 months) with the results for any children less well matched on age.

4. The authors propose to undertake the case-control study with 'ever vaccinated' as the risk factor. It might be as well to plan analyses in which exposure is taken as vaccination within fixed periods prior to the index date, as I would suspect that the power of the study might otherwise be very low. The design assumption of 85% vaccination coverage is unrealistically low in the cohort born after 1987: the coverage by 2 years is more likely to be closer to 95%.

We agree with the suggestion to undertake analyses in which exposure is taken as vaccination within fixed periods prior to the index date: in fact this is already included in the analyses section ("Two analyses will be carried out. The first will estimate the odds ratio for the development of symptoms in specific time periods after vaccination with MMR").

Minor comments

1. p7 lines 11-13. In the Taylor et al study, risk periods of up to 2 years were investigated, not just 6 months.

We agree and have altered this.

2. p7 line 16. Delete journal name and date from reference 21: the paper has not yet been accepted for publication.

We agree: particularly as the referee is the lead author of the paper in question!

3. *p14 lines 13-14. The authors state that 'there are good reasons to expect the vaccine data to be complete'. Will this be checked?*

Because the data is entirely anonymous, we will not be able to check vaccination records against health authority child health records (the method used by Taylor et al[1]). We will be collecting information about vaccination in the parental questionnaire.

4. *p14 line -1. Will separate analyses be undertaken for the two sets of controls? If so, how will the authors interpret discordant results?*

Yes: this is now fully explained in the paper.

5. *p15 lines 9-11. The case series method does not necessarily require a defined post-vaccination risk period, as shown by reference 21, and sidesteps the problem of individual-level confounding.*

We agree with the first point and have amended the paper accordingly. We have some disagreement with the second point. However we do agree that the case-series is a powerful method with great scope to reduce confounding: hence our decision to use it in this study.

Review by Charles Hoff Comments

There are, however, some minor syntactical issues which should be addressed by the authors or editorial staff. If some of these are due to differences in practice between U.S. and U.K. English, then ignore them.

- a) p. 9., line 4: there should be an 'of' between the words 'copies' and 'letters',*
- b) p 10., line 5: dangling prep.; I suggest saying "practice with whom participants are registered.",*
- c) p10, line 10: ditto; I suggest saying "child with whom they are matched.",*
- d) p. 11, line 3: there should be a comma between 'addition' and 'the',*
- e) p. 11, line 6; dangling prep.; I suggest saying "with and those without reported regression.",*
- f) p13, line 1: "exclusion to at" should be "exclusion of at",*
- g) p. 13, line 7: there should be a comma after 'e.g.',*
- h) p. 14, line 5: ditto comma after 'analyses',*
- i) p. 14, line 10: "has a two" should be "has two",*
- j) p .14, lines 11, 12, 13: there should be commas after 'Firstly', 'Secondly', and 'Finally'
— respectively,*
- k) p. 17, line 12 and 20: ditto after 'particular' and 'However',*
- l) throughout the ms. 'case series' should be 'case-series' and 'case control' should be 'case-control'.*
- m) there are possibly more places where commas are required which I missed.*

Some of the items in the above list were simple errors. Apologies: we have corrected these. We have not changed issues of style: we leave these to the editorial staff.

2. The proposed use of conditional logistic regression seems appropriate. However, since there is the issue of age and elapsed time from vaccination to when the patient's status is examined, survival analytic approaches would also seem appropriate (e.g., stepwise Cox regression). I believe that this would add more 'robustness' in clarifying associations between vaccination and outcome. My logic here is that there is the possibility that some younger examined patients with short elapsed times might yet develop autism after the study is completed. I encourage the authors to discuss this with an epidemiologic-oriented biostatistician. I have not checked this in the cited papers but I will bet 'dollars to doughnuts' that none of the studies cited used this approach. This might give the authors 'a leg up' in their investigation!!!

There are two concerns implicit in this comment. The first is what are the consequences of a control developing autism. As autism is very rare, this is extremely unlikely and if it happened would not distort our result to any substantial degree. Second, that cases and controls are comparable in the time they were at risk of developing autism. As cases and controls are matched for date of birth, they will have comparable time at risk. We therefore do not think the recommendation is necessary and the existing design of the study is adequate to address these concerns.

1. Taylor B, Miller E, Farrington CP, Petropoulos MC, Favot-Mayaud I, Li J *et al.*: **Autism and measles, mumps, and rubella vaccine: no epidemiological evidence for a causal association.** *Lancet* 1999, **353**: 2026-2029.