

Meta-analytic approaches to determine gender differences in the age-incidence characteristics of schizophrenia and related psychoses

DAN JACKSON, JAMES KIRKBRIDE, TIM CROUDACE, CRAIG MORGAN, JANE BOYDELL,
ANTONIA ERRAZURIZ, ROBIN M. MURRAY & PETER B. JONES

MRC Biostatistics Unit, Robinson Way, Cambridge, UK

Key words

fractional polynomial, incidence rates, meta-regression

Correspondence

Dan Jackson, MRC Biostatistics Unit, Robinson Way, Cambridge, CB2 2SR, UK.
Telephone (+44) (0)1223 330386
Email: daniel.jackson@mrc-bsu.cam.ac.uk

Abstract

A recent systematic review and meta-analysis of the incidence and prevalence of schizophrenia and other psychoses in England investigated the variation in the rates of psychotic disorders. However, some of the questions of interest, and the data collected to answer these, could not be adequately addressed using established meta-analysis techniques. We developed a novel statistical method, which makes combined use of fractional polynomials and meta-regression. This was used to quantify the evidence of gender differences and a secondary peak onset in women, where the outcome of interest is the incidence of schizophrenia. Statistically significant and epidemiologically important effects were obtained using our methods. Our analysis is based on data from four studies that provide 50 incidence rates, stratified by age and gender. We describe several variations of our method, in particular those that might be used where more data is available, and provide guidance for assessing the model fit. *Copyright © 2013 John Wiley & Sons, Ltd.*

Introduction

A recent systematic review and meta-analysis of the incidence and prevalence of schizophrenia and other psychoses in England (Kirkbride *et al.*, 2011, 2012) used a range of meta-analytic techniques, including random effects modelling, meta-regression and multivariate meta-analysis, to inspect variation in the incidence and prevalence of psychotic disorders. However, some of the questions of interest, and the data collected to answer these, could not be adequately addressed using established meta-analysis techniques. The aim of this paper is to present the methodological approach that we developed

to overcome these difficulties, and provide details of the results. Although the new methodology was applied to several different psychotic syndromes, here we will describe and present results for only one outcome measure, the incidence of schizophrenia.

Gender differences in the incidence and prevalence of schizophrenia are well established from individual surveys (Angermeyer and Kuhn, 1988; Flor-Henry, 1985; Kirkbride *et al.*, 2006). Typically, incidence peaks in the early twenties for men and a few years later for women. Until the mid-thirties, rates are typically estimated to be approximately 1.5 to two times greater in men

than women. After this time, rates decline for both sexes, with a narrowing sex ratio, until the mid-forties when there is a smaller secondary peak for women. Such variation may have important implications for understanding the underlying aetiology of the disorder (Hafner *et al.*, 1993), particularly with regard to a putative neuroprotective role for estrogen prior to menopause or the biopsychosocial impact of such experiences. We wished to investigate these differences using the combined precision and power available from a systematic review. In particular, we wished to investigate whether there was evidence of an elevated incidence rate in men compared to women, and later in life for women compared with younger women.

Methodology

Four suitable data sources were identified from the systematic review (Kirkbride *et al.*, 2012). Three of these sources came from published studies (Kirkbride *et al.*, 2006; Brewin *et al.*, 1997; Goldacre *et al.*, 1994), with the final dataset coming from data unpublished at the time of the report (since published; Reay *et al.*, 2010). These sources provided data which motivated the development of our approach to studying sex and age-related variations in incidence in a more flexible manner. Here, each primary study provided the natural logarithm of the raw (unadjusted) incidence rate and its standard error (s.e.) for the study population, for men and women separately, stratified by age group (typically deciles of age). The logarithm of the incidence rate is a standard measure used in meta-analysis for which a normal approximation is conventionally used (Sutton *et al.*, 2000). Eligible studies, however, did not all use the same age strata when reporting incidence rates so age mid-points were used when modelling the data (see Tables 1 and 2).

The data required for analysis are shown in Table 2. From such data we wanted to perform a single meta-analysis to obtain simultaneous statistical inferences to address the following questions:

- (1) Is there evidence of a sex difference in rates?
- (2) Is there evidence of changed rates for women during mid-life?
- (3) Are rates for men and women more similar in later life than at younger ages?

Our statistical methods were variations and novel applications of new methodology introduced in a recent paper by Rota *et al.* (2010). Our trends were over the population's age, whereas those of Rota *et al.* (2010) were over the drug's exposure level, but the statistical issues are

similar because, in both cases, flexible trend modelling in the context of meta-analysis was desired. In our application, given prior knowledge in this field, a linear model for the effect of age was inappropriate. Furthermore a simplistic and incorrect assumption e.g. that of linearity, here might adversely influence our inferences for the effects described in the introduction that are of real interest. Instead, we wanted a flexible model of the effects of age, using as few parameters as possible, because our analysis was based on only 50 rates (Table 2). Because the age distribution of schizophrenia rates is often reported to exhibit a secondary peak onset (SPO) in women during their mid-forties, we sought to examine evidence for such variation around this period by explicitly modelling this in our meta-regression. Here we defined 45 years to be the mid-point for the SPO.

We decided, a priori, to avoid a two stage approach such as that of Rota *et al.* (2010) because this involves estimating the between-study variance structure using a very small number of studies. For our data, with a relatively small number of contributing studies, it is clear that between-study variance parameters estimates will be very imprecise. Two-stage approaches are sometimes advocated in the context of meta-analysis with large datasets (The Fibrinogen Studies Collaboration, 2009), but with just 50 rates to model this is not tenable. Estimation in the second stage of Rota *et al.*'s (2010) approach depends directly on this estimated variance structure and problems with using random effects models in meta-analyses, even in the univariate case, when there is a small number of studies is now generally appreciated (e.g. Jackson and Bowden, 2009). We further desired a simple and computationally straightforward approach because we intended to use the methodology routinely for a variety of outcome measures (i.e. different psychotic disorders) relevant to the main systematic review (Kirkbride *et al.*, 2012).

Hence we used the model:

$$Y_{it} = \alpha_i + \beta_1 x_{it}^{p_1} + \beta_2 x_{it}^{p_2} + \delta m_{it} + \phi n_{it} + \varepsilon_{it} \quad (1)$$

where Y_{it} was the t th logarithm of the crude (i.e. unadjusted) incidence rate from the i th study and x_{it} was the midpoint of the corresponding age stratum. The terms $\beta_1 x_{it}^{p_1} + \beta_2 x_{it}^{p_2}$ were the fractional polynomial used to model age trends in the data (see later). The variables m_{it} and n_{it} were indicators for the corresponding incidence rate for men, and women with an age-stratum midpoint of at least 45 years, respectively. Hence δ denoted one parameter of primary interest, a gender effect, and ϕ denoted the other, an effect (raised rate) in women after

Table 1. Data sources: study overview and methodology

Authors	Setting	Case ascertainment period	Age range (years) (schizophrenia)	Sample size	Design ²	Diagnosis ³	Methodology overview
Brewin <i>et al.</i>	Nottingham	1992–1994 (2 years)	16–64	57	Case register	CD-10	All cases presenting with a first episode of psychosis in a given catchment area identified over two-year period. Exclusion criteria applied for subjects out of age range or study area, or with an organic basis to disorder. Denominator estimated from census data
Goldacre <i>et al.</i>	Oxfordshire	1975–1986 (11 years)	15–64	2,024	First contact	CD-7, ICD-8, ICD-9	Analysis of linked records in Oxfordshire for all people in contact with specialist psychiatric services As Brewin <i>et al.</i>
Kirkbride <i>et al.</i>	SE London, Nottinghamshire & Bristol	1997–1999 (2 years)	16–64	209	First contact	CD-10	
Reay <i>et al.</i>	Northumberland	1998–2005 (7 years)	16–69	60	First contact	CD-10	Observational database was set up on all patients presenting with FEP to secondary services in catchment area. Denominator obtained from census

¹Unpublished at the time of systematic review. For published data, see Reay *et al.* (2010).

²Case register: all people diagnosed with a psychotic disorder for the first time registered on a disease surveillance database. First contact: study designed to identify all potential service bases where people with a psychotic disorder may present for the first time. Service bases screened regularly and cases usually independently verified by standardized psychiatrist diagnosis.

³ICD: International Classification of Disease (and version).

Table 2. The natural logarithm of 50 incidence rates, stratified by gender and age group, from the four studies described in Table 1

Study	log(IR)	s.e.	Gender	Age midpoint	Study	log(IR)	s.e.	Gender	Age midpoint
1	1.79	0.79	Male	17.5	3	0.36	1	Male	47
1	3.18	0.2	Male	24.5	3	1.44	0.58	Male	52
1	2.2	0.38	Male	34.5	3	1.27	0.71	Male	57
1	0	0.91	Male	44.5	3	2.09	0.5	Male	62
1	1.1	0.66	Male	54.5	3	2.3	0.41	Female	17.5
1	1.79	0.68	Female	17.5	3	2.24	0.33	Female	22
1	2.2	0.37	Female	24.5	3	2.41	0.29	Female	27
1	0.69	0.64	Female	34.5	3	2.37	0.29	Female	32
1	1.61	0.65	Female	44.5	3	2.03	0.35	Female	37
1	0.69	0.87	Female	54.5	3	1.95	0.41	Female	42
1	1.1	0.89	Female	62	3	2.12	0.41	Female	47
2	2.8	0.09	Male	19.5	3	1.04	0.71	Female	52
2	2.82	0.1	Male	29.5	3	1.64	0.58	Female	57
2	2.24	0.15	Male	39.5	3	0.66	1	Female	62
2	1.69	0.16	Male	54.5	4	2.36	0.38	Male	17.5
2	2.16	0.13	Female	19.5	4	2.98	0.22	Male	24.5
2	2.2	0.14	Female	29.5	4	2.17	0.28	Male	34.5
2	2.12	0.16	Female	39.5	4	0.23	0.71	Male	44.5
2	1.79	0.15	Female	54.5	4	0.66	0.58	Male	54.5
3	3.48	0.22	Male	17.5	4	0.45	1	Female	17.5
3	3.67	0.16	Male	22	4	1.36	0.5	Female	24.5
3	3.17	0.2	Male	27	4	-0.44	1	Female	34.5
3	3.08	0.2	Male	32	4	0.92	0.5	Female	44.5
3	2.95	0.22	Male	37	4	0.65	0.58	Female	54.5
3	1.97	0.41	Male	42	4	-0.17	1	Female	64.5

the SPO. In situations where the SPO is thought to occur at alternative ages in different studies, n_{ii} would be defined as an indicator for women with an age-stratum midpoint that exceeds the study specific SPO. Although we use the term “effect” when describing our parameters, it is important to be clear that our data are observational and no causal associations are implied by this term.

Positive δ would indicate that women had lower incidence rates than men until the SPO. Positive ϕ would indicate that there was an elevated risk of schizophrenia in women after the SPO compared with women immediately before this. However, a positive ϕ would *not* imply that older women had higher incidence rates than younger women; it is known that the peak age of onset for schizophrenia is most common in early adulthood. This is because both study and age were included in Equation 1, so estimates of δ and ϕ were adjusted for these. In particular, the fractional polynomial in Equation 1 modelled the effect of age. This can, for example, model incidence rates that are strongly decreasing for women as their age increases, so that older women have lower rates than

younger women, despite a positive ϕ and hence a step change increase in the incidence rate around the SPO. This can be assessed by examining the model’s predictions, as shown in Figures 1 and 2. When drawing inferences we focused on the estimates of δ and ϕ , because they address the questions that particularly interested us. However, depending on the questions of interest, any aspect of the model fit could be used for interpretation, including the α_i parameters. In addition to enabling us to make the inferences we require, Equation 1 provides a way to explain why the observed rates differ.

The α_i were fixed effects, stratified by study, and model the inevitable association between incidence rates from the same study (e.g. Salanti *et al.*, 2008). Since the pooled or study specific rates were of little interest, with δ and ϕ being primary, we chose to model the study effects in the simplest and most convenient way for our purposes. Although the α_i allowed for between-study heterogeneity, in the sense that different studies were permitted to have elevated or reduced rates, the model assumes common fractional polynomials, gender effects δ and SPO effects ϕ

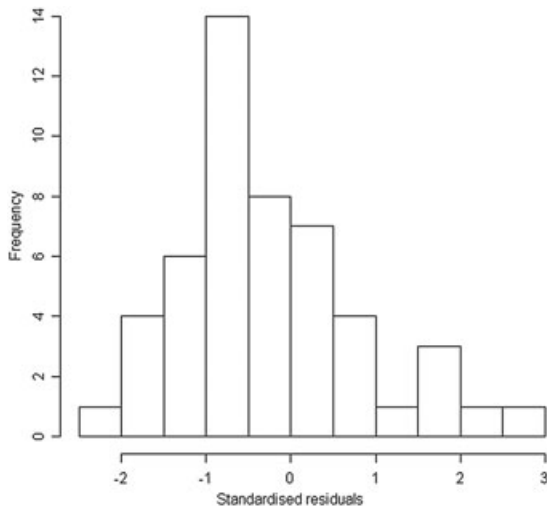


Figure 1. Histogram of 50 standardized residuals from fitting Equation 1.

in women across studies. If these assumptions are not true then the resulting standard errors will, in general, be too small. We return to this issue in the discussion. However our fixed effect approach has the advantage

that the model is easier to interpret because we do not interpret regression coefficients as average values, as we do in random effects modelling.

The term ε_{it} was the normally distributed error, with variance taken from study reports. These variances were regarded as fixed and known, a priori, as is conventional in meta-analysis (e.g. Biggerstaff and Tweedie, 1997; Hardy and Thompson, 1996). Their square-root, the standard errors (s.e. values), are shown in Table 2.

Fractional polynomials

Fractional polynomials (Royston and Altman, 1994) are an established way to parsimoniously model trends. We adopt the convention, also used by Rota *et al.* (2010), of selecting p_1 and p_2 from a predefined set $P = \{-2, -1, -0.5, 0, 0.5, 1, 2, 3\}$, a set of 36 trend functions, including U shaped and J-shaped relations. Powers of $p_1 = 0$ and $p_2 = 0$ were taken to indicate taking the natural logarithm. If $p_1 = p_2$ then $x_{it}^{p_2}$ is replaced by $x_{it}^{p_2} \log(x_{it})$ in the fractional polynomial (Royston and Altman, 1994, Rota *et al.*, 2010), so that the regression parameters are identifiable in such instances. We follow Rota *et al.*

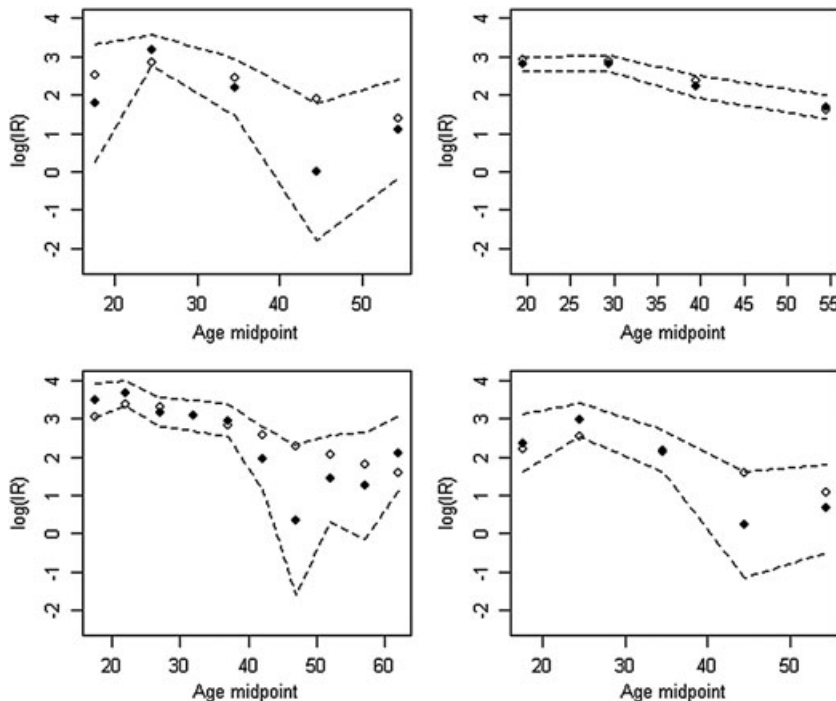


Figure 2. Male observed (solid points) and fitted (hollow points) log incidence rates in each of the four studies. Ninety-five per cent confidence regions, obtained from the rates and standard errors shown in Table 2, are also shown, where the upper and lower bounds of the intervals are connected using dashed lines to aid interpretation.

(2010), in choosing p_1 and p_2 as the best fitting fractional polynomial (minimum weighted sum of squared residuals, where the weights are those used when fitting Equation 1) over all studies and basing all inference on the resulting model, ignoring the uncertainty in the form of the fractional polynomial. This followed the same principle as Rota *et al.* (2010), who determined the best fitting fractional polynomial using the Akaike information criterion (AIC).

Meta-regression

Meta-regression (Thompson and Sharp, 1999) is a well established technique. Once the form of the fractional polynomial has been chosen, Equation 1 has the same form as a standard fixed effects meta-regression. Furthermore it is a simple extension of Equation 1 of Thompson and Sharp (1999), which included just a single covariate. Hence, once p_1 and p_2 have been determined, Equation 1 can be fitted using a standard linear regression, where the weights are given by the reciprocal of the variances of ε_{it} . The correct standard errors of the regression coefficients were, however, obtained by dividing those given by the square root of the reported mean square error (as explained by Thompson and Sharp, 1999). Our Equation 1 was therefore very easy to fit in any statistical software for standard linear models.

Implementation

R software was used to implement the method. First, each possible combination of powers p_1 and p_2 was used in the fractional polynomial when fitting Equation 1. The model that minimized the resulting weighted sum of squares was taken as the best fitting model. Equation 1 was fitted as a standard (weighted) linear regression model using the *lm* command, with the weights defined as the reciprocal of the variances and where the response variable was the logarithm of the incidence rate.

The best fitting model was then used for inference, where the standard errors were corrected as explained earlier.

Results

The best fitting fractional polynomial for the incidence of schizophrenia data was $p_1 = -1$ and $p_2 = -1$. This provided a weighted sum of squared residuals of 63.39. Fitting Equation 1 with $p_1 = p_2 = -1$ required estimating eight parameters (four α_i terms, β_1 , β_2 , δ and ϕ) from 50 rates, emphasizing the case for our parsimonious model.

The resulting estimates were $\hat{\delta} = 0.69$ (s.e.: 0.08) and $\hat{\phi} = 0.71$ (s.e.: 0.18). Confidence intervals can be obtained as parameter estimates plus and minus 1.96 standard errors and similarly p -values can be calculated from these results.

Both estimates resulted in the inferences we anticipated with the interpretation of these coefficients as follows: the estimate of δ was positive and highly statistically significant, indicating that the incidence of schizophrenia was greater for men than women prior to the SPO; the corresponding point estimate of the incidence ratio was approximately two. The estimate of ϕ is of similar magnitude and was also statistically significant, but this parameter is less precisely estimated. This relative imprecision seems reasonable, because only nine of the 50 rates are taken from women after the SPO, so this effect must be harder to identify. Nevertheless, this analysis suggests that the incidence ratio for women just before, and just after, the SPO is also around two.

The difference between the log incidence rate of schizophrenia in men and women after the SPO is $\delta - \phi$. The estimate of this difference was $\hat{\delta} - \hat{\phi} = -0.02$, and the standard error of this estimated difference was

$$\sqrt{\text{Var}(\hat{\delta}) + \text{Var}(\hat{\phi}) - 2\text{Cov}(\hat{\delta}, \hat{\phi})}$$

The correlation of the estimates of δ and ϕ was obtained when fitting Equation 1 as a weighted linear regression as explained earlier, and so their covariance and hence the standard error of the estimated difference was calculated as 0.17. The data did not provide enough information to draw any firm conclusions about the direction of the difference between incidence rates of schizophrenia for men and women in the oldest age groups, but it seems reasonable to infer that these are more similar than those for younger people.

A considerable advantage of our methods was that, since they were based on standard linear models, the examination of model diagnostics such as leverage, influence and residuals are straightforward. Such diagnostics are also considered important in meta-analysis (Viechtbauer and Cheung, 2010). We calculated (approximately) standardized residuals as the ratio of the raw residuals and the studies' standard errors in Table 2. If Equation 1 fits well then these standardized residuals should resemble a standard normal distribution. A histogram of the 50 standardized residuals is shown in Figure 1. The residuals were slightly skewed, but Figure 1 reassures us that the model fit was quite reasonable for these data.

Plots of the observed, and fitted, rates were shown separately for men and women in each of the four studies in Figures 2 and 3, respectively. A comparison of these

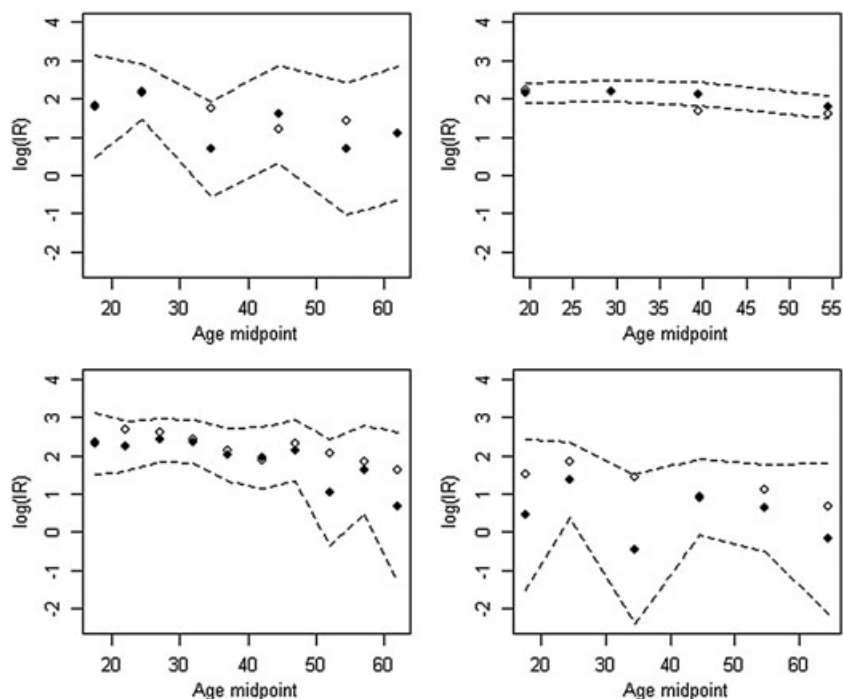


Figure 3. Female observed (solid points) and fitted (hollow points) log incidence rates in each of the four studies. Ninety-five per cent confidence regions, obtained from the rates and standard errors shown in Table 2, are also shown, where the upper and lower bounds of the intervals are connected using dashed lines to aid interpretation.

plots supported the formal statistical inference that rates were higher in young men than young women. Further, this difference appeared to be reduced in the older population. Although the model captured the general trends in the data quite well, it did not capture the apparent fall in the incidence rate for men in their forties, and subsequent increase in their fifties, that three of the four studies were suggestive of (Figure 2). If this is a real phenomenon and we wish to describe this, further modelling is required. It is perhaps unfortunate that these unexpected falls occurred at around the age where the model provides a step change for women. The apparent discrepancies between the observed and fitted rates for men in their forties are probably due to a combination of the problems associated with modelling real data and our relatively simple model.

The overall impression of Figures 1–3 is that our modelling has performed well, but further refinement may be possible.

Discussion

We have developed a novel methodology in order to flexibly model variation in incidence rates for schizophrenia.

Specifically, this involved combining fractional polynomials with meta-regression. Importantly our method avoided the difficulties associated with attempting to estimate any between-study variance parameters. However, it achieved this by assuming common fractional polynomials, gender, and SPO in women, effects across studies. These are quite strong assumptions but the resulting model adequately describes all the datasets in our review.

The simplest way to relax these assumptions is to adopt a two stage approach and fit Equation 1 to each study in turn and then combine the resulting estimates in a conventional random effects meta-analysis at a second stage. With only four studies, the resulting estimation of the between-study variation here is very imprecise but for examples with more studies this is a more viable solution to allow heterogeneous effects of interest. Despite our reservations in doing so, we fitted univariate random effects models to the four study specific estimates of δ and ϕ in order to explore the implications of this alternative procedure. The resulting random effects model for ϕ collapsed to a fixed effects model and similar inferences were obtained, but the random effects model for δ provided a large I^2 statistic (Higgins and Thompson, 2002) of 57%, despite the lack of a strongly statistically significant hypothesis test for

the presence of heterogeneity (p -value = 0.07). Although a broadly similar point estimate of δ was produced by this random effects analysis, a much larger corresponding standard error of 0.15 was obtained. This is due to the large, but very imprecise, estimate of the between-study variance. This finding reinforces our statement that standard errors from our procedure will in general be too small if the common study effects model (Equation 1) assumes are not true and diagnostics, such as Figure 1, are essential to assess the evidence of any over-dispersion.

If sufficient data were available we would suggest pooling the estimates of δ and ϕ in a bivariate random effects meta-analysis at a second stage in a two stage approach. We leave the question concerning how much data we need, in order to favour a multivariate meta-analysis to the proposed univariate meta-regression based procedure, as an open question for further research. If there were very many studies providing information on the same age groups then we anticipate that multivariate random effects meta-analysis could prove useful for analysing heterogeneous data where the incidence rates, rather than the estimates of δ and ϕ , are the outcomes. In fact bivariate meta-analysis was used in the systematic review, to meta-analyse epidemiological studies that gave information by sex but whose published results did not offer rates stratified by age.

In some instances more sophisticated one-stage random effects modelling may be desirable and plots such as Figures 1–3 will be essential in order to determine whether this is needed. In situations where the observed rates appear to be more variable than predicted by the model, one straightforward way to proceed would be to incorporate a random effect into the resulting meta-regressions in the usual way (Thompson and Sharp, 1999). However this standard methodology assumes that each study provides a single rate, so that the random effects are independent, but here it is plausible that any random effects within-studies may be correlated. For example, it could be that the unknown additional factors that result in between-study variation for a particular study provide increased rates for older people, in which case the male and female random effects for particular age groups would be correlated. Since both the point estimates and the confidence intervals depend on both the modelling and the resulting estimates of the random effect structure, any particular one-stage random effects model fit, including those that model heterogeneity in the parameters of particular interest, could be hard to defend for the type of dataset that we have analysed.

Another pragmatic approach to dealing with any apparent over-dispersion is to use Thompson and Sharp's (1999) non-hierarchical based multiplicative

over-dispersion parameter (their model 2) when fitting the meta-regressions, although Thompson and Sharp (1999) do not recommend this particular approach in practice when conducting more conventional meta-regressions. This results in the same point estimates as our procedure but standard errors are not divided by the square root of mean square error. This increases the standard error of estimates of δ and ϕ by around 23% in our example, but statistically significant findings are still obtained. More sophisticated hierarchical modelling may be desirable in some instances but there are limits to what may be achieved when the available data comprise of just 50 rates from the available literature (Kirkbride *et al.*, 2012). Rather than devoting any additional effort to the modelling of the random effects, it may instead be preferable to focus on the form of the regression in Equation 1, which assumes that rates vary continuously for men, but that there is a step change at SPO for women.

Quantifying and measuring the impact of the between-study heterogeneity is often considered important in meta-analysis (Higgins and Thompson, 2002) and our approach somewhat complicates the assessment of this aspect of primary study data, because rather than providing a between-study variance parameter on which measures of heterogeneity (Higgins and Thompson, 2002) and inference (Biggerstaff and Jackson, 2008) can be based, our approaches uses a fixed effects approach. There is clear evidence that the estimated study specific effects α_i differ in our sample but we have resisted the temptation to test for homogeneity, $H_0: \alpha_i = \alpha$, for all α_i , as usually recommended (Higgins and Thompson, 2002).

In general we found between-study variation was very large in our systematic review; where standard random effects meta-analysis was used, I^2 statistics (Higgins and Thompson, 2002) in excess of 90% were not uncommon which we interpret as being indicative of very considerable heterogeneity. Partly because of these very large degrees of heterogeneity, and also because we were pooling rates from observational studies, when we presented pooled results from random effects meta-analyses in our systematic review we did so as *descriptive* rather than inferential statistics. Although we have performed formal inference for using our two parameters of interest here, we wish to make it clear that although our model adequately describes the data, all the usual cautions and caveats when interpreting statistical findings from observational studies should still be exercised here.

Our methodology was equally successful at detecting and describing the covariate effects of gender and SPO in women in many other study sets identified in the review, including for other epidemiological outcome measures.

By including these covariate effects of interest, in addition to the non-linear trend modelled by the fractional polynomial, our method provides a way of extending the type of approach introduced by Rota *et al.* (2010). Our methods made salient the features of this data that would be expected given existing results in the literature, while highlighting new features that were not necessarily obvious from simple visual inspection of the data (Table 2).

We used the midpoint of the age groups as covariates in our fractional polynomials and we recognize that the categorization of continuous variables can present genuine statistical issues (Altman *et al.*, 1994). Furthermore we took the best fitting fractional polynomial and based all formal statistical inference on this, and we also recognize that this strategy is not without its limitations. Hence we provide the relevant data in Table 2 in the hope that it might continue to motivate methodological development on alternatives. R code is available from the first author on request to perform the type of analysis

used and to facilitate the development of this type of statistical technique.

Acknowledgements

This is an independent report commissioned and funded by the Policy Research Programme in the Department of Health. The views expressed are not necessarily those of the department. Dan Jackson was supported by MRC grant U.1052.00.006.

James Kirkbride was supported by a Sir Henry Wellcome Research Fellowship from the Wellcome Trust (Grant code: WT085540). Tim Croudace was supported in part by a DoH/NIHR Career Scientist Award in Public Health. The study was conducted within the NIHR Collaboration for Leadership in Applied Health Research and Care (CLAHRC) for Cambridgeshire and Peterborough.

The authors thank Ian R White for his constructive comments and advice on earlier versions of this paper.

Declaration of interest statement

The authors have no competing interests.

References

- Altman D.G., Lausen B., Sauerbrei W. (1994) Commentary: dangers of using "optimal" cutpoints in the evaluation of prognostic factors. *Journal of the National Cancer Institute*, **86**(11), 829–835.
- Angermeyer M.C., Kuhn L. (1988) Gender differences in age at onset of schizophrenia. An overview. *European Archives of Psychiatry and Neurological Sciences*, **237**(6), 351–364.
- Biggerstaff B., Jackson D. (2008) The exact distribution of Cochran's heterogeneity statistic in one-way random effects meta-analysis. *Statistics in Medicine*, **29**(29), 6093–6110, DOI: 10.1002/sim.3428
- Biggerstaff B., Tweedie R. (1997) Incorporating variability in estimates of heterogeneity in the random effects model in meta-analysis. *Statistics in Medicine*, **16**(7), 753–768.
- Brewin J., Cantwell R., Dalkin T., Fox R., Medley I., Glazebrook C., Kwiecinski R., Harrison G. (1997) Incidence of schizophrenia in Nottingham – a comparison of two cohorts, 1978–80 and 1992–94. *The British Journal of Psychiatry*, **171**(2), 140–144.
- Flor-Henry P. (1985) Schizophrenia: sex differences. *Canadian Journal of Psychiatry*, **30**(5), 319–322.
- Goldacre M., Shiwach R., Yeates D. (1994) Estimating incidence and prevalence of treated psychiatric disorders from routine statistics: the example of schizophrenia in Oxfordshire. *Journal of Epidemiology and Community Health*, **48**(3), 318–322.
- Hafner H., Riecher-Rossler A., An Der Heiden W., Maurer K., Fatkenheuer B., Loffler W. (1993) Generating and testing a causal explanation of the gender difference in age at first onset of schizophrenia. *Psychological Medicine*, **23**(4), 925–940.
- Hardy R.J., Thompson S.G. (1996) A likelihood approach to meta-analysis with random effects. *Statistics in Medicine*, **15**(6), 619–629.
- Higgins J.P.T., Thompson S.G. (2002) Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, **21**(11), 1539–1558, DOI: 10.1002/sim.1186
- Jackson D., Bowden J. (2009) A re-evaluation of the 'quantile approximation method' for random effects meta-analysis. *Statistics in Medicine*, **28**(2), 338–348, DOI: 10.1002/sim.3487
- Kirkbride J.B., Errazuriz A., Croudace T.J., Morgan C., Jackson D., McCrone P., Murray R.M., Jones P.B. (2011) Systematic Review of the Incidence and Prevalence of Schizophrenia and Other Psychoses in England. Conducted for the Department of Health Policy Research Programme. <http://www.psychiatry.cam.ac.uk/epicentre/review/> [13 March 2012].
- Kirkbride J.B., Errazuriz A., Croudace T.J., Morgan C., Jackson D., Boydell J., Murray R.M., Jones P.B. (2012) Incidence of schizophrenia and other psychoses in England, 1950–2009: a systematic review and meta analyses. *PLoS One*, **7**(3), e31660.
- Kirkbride J.B., Fearon P., Morgan C., Dazzan P., Morgan K., Tarrant J., Lloyd T., Holloway J., Hutchinson G., Leff J.P., Mallett R.M., Harrison G.L., Murray R.M., Jones P.B. (2006) Heterogeneity in incidence rates of schizophrenia and other psychotic syndromes: findings from the 3-Center AESOP study. *Archives of General Psychiatry*, **63**(3), 250–258.
- Reay R., Mitford E., McCabe K., Paxton R., Turkington D. (2010) Incidence and diagnostic diversity in first-episode psychosis. *Acta Psychiatrica Scandinavica*, **121**(4), 315–319.
- Rota M., Bellocco R., Scotti L., Tramacere I., Jenab M., Corrao G., La Vecchia C., Boffetta P., Bagnardi V. (2010) Random-effects meta-regression models for studying nonlinear dose–response relationship, with an application to alcohol and esophageal squamous cell carcinoma. *Statistics in Medicine*, **29**(26), 2679–2687, DOI: 10.1002/sim.4041
- Royston P., Altman D.G. (1994) Regression using fractional polynomials of continuous covariates: parsimonious parametric modelling (with Discussion). *Applied Statistics*, **43**(3), 429–467.
- Salanti G., Higgins J.P., Ades A., Ioannidis J.P. (2008) Evaluation of networks of randomized trials. *Statistical Methods in Medical Research*, **17**(3), 279–301.

- Sutton A.J., Abrams K.R., Jones D.R., Sheldon T.A., Song F.. (2000) Methods for Meta-analysis in Medical Research, Chichester: John Wiley & Sons.
- The Fibrinogen Studies Collaboration. (2009) Systematically missing confounders in individual participant data meta-analysis of observational cohort studies. *Statistics in Medicine*, **28**(8), 1218–1237, DOI: 10.1002/sim.3540
- Thompson S.G., Sharp S.J. (1999) Explaining heterogeneity in meta-analysis: a comparison of methods. *Statistics in Medicine*, **18**(20), 2693–2708.
- Viechtbauer W., Cheung M.W.-L. (2010) Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods*, **1**(2), 112–125, DOI: 10.1002/jrsm.11