Practice of Epidemiology

Risk of Pelvic Inflammatory Disease Following *Chlamydia trachomatis* Infection: Analysis of Prospective Studies With a Multistate Model

Malcolm J. Price*, A. E. Ades, Daniela De Angelis, Nicky J. Welton, John Macleod, Kate Soldan, Ian Simms, Katy Turner, and Paddy J. Horner

* Correspondence to Dr. Malcolm J. Price, School of Health and Population Sciences, Public Health Building, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom (e-mail: m.price.2@bham.ac.uk).

Initially submitted May 22, 2012; accepted for publication December 18, 2012.

Our objective in this study was to estimate the probability that a *Chlamydia trachomatis* (CT) infection will cause an episode of clinical pelvic inflammatory disease (PID) and the reduction in such episodes among women with CT that could be achieved by annual screening. We reappraised evidence from randomized controlled trials of screening and controlled observational studies that followed untreated CT-infected and -uninfected women to measure the development of PID. Data from these studies were synthesized using a continuous-time Markov model which takes into account the competing risk of spontaneous clearance of CT. Using a 2-step piecewise homogenous Markov model that accounts for the distinction between prevalent and incident infections, we investigated the possibility that the rate of PID due to CT is greater during the period immediately following infection. The available data were compatible with both the homogenous and piecewise homogenous models. Given a homogenous model, the probability that a CT episode will cause clinical PID was 0.16 (95% credible interval (CrI): 0.06, 0.25), and annual screening would prevent 61% (95% CrI: 55, 67) of CT-related PID in women who became infected with CT. Assuming a piecewise homogenous model with a higher rate during the first 60 days, corresponding results were 0.16 (95% CrI: 0.07, 0.26) and 55% (95% CrI: 32, 72), respectively.

Bayesian analysis; causal effect; *Chlamydia trachomatis*; Markov model; mass screening; meta-analysis; pelvic inflammatory disease; prospective studies

Abbreviations: CT, *Chlamydia trachomatis*; CT+, CT-infected; CT−, CT-uninfected; PID, pelvic inflammatory disease; POPI, Prevention of Pelvic Infection; STI, sexually transmitted infection.

Pelvic inflammatory disease (PID) is a leading cause of both tubal factor infertility and ectopic pregnancy (1, 2). PID is a clinical diagnosis typically indicated by lower abdominal pain, in association with local tenderness upon bimanual vaginal examination (3). Infection with *Chlamydia trachomatis* (CT) is an important risk factor for the development of PID, and knowledge of the role of CT is critical for assessing the efficacy and cost-effectiveness of CT screening (4, 5). PID is also caused by other infections of the female reproductive tract, including other sexually transmitted infections (STIs) and infections introduced during surgery, abortion, or parturition; these other risk factors may be unaffected or indirectly affected by CT screening and cause a background level of nonchlamydial PID (3). A range of observational studies have followed untreated CT-infected women prospectively and have observed the numbers of women developing PID (6–10); in some studies the infection, presumably symptomatic, has been diagnosed in a clinic (6–8, 10), and in 1 study it was identified by screening asymptomatic cases (9). There have also been controlled trials randomizing large numbers of women to screening and treatment or to no screening (11–13). However, there is a remarkable lack of consensus on the risk of PID attributable to an episode of CT, in spite of a major authoritative review (1). Still more surprising is the lack of consensus on what studies should be used to estimate this risk.

In the United Kingdom, one recent cost-effectiveness analysis of screening (5) estimated the risk of PID following...
CT from a Swedish data linkage study (14), and another based an estimate on retrospective data (15). Neither used the evidence from randomized controlled trials, although this would be regarded as the most reliable evidence of causal effects.

Programs designed to screen for and treat CT have 2 purposes: to reduce transmission to partners and thereby lower prevalence and the number of future infections, and to prevent PID and further sequelae in the women found to be infected (4, 16). Untreated CT infection usually clears within 1.5 years (17–20), so women who are diagnosed through screening have already “survived” a period of infection without developing PID. Randomized controlled trials of screening and treatment versus no screening can provide direct information on the proportion of CT-caused PID that can be prevented in a population with prevalent infection (13, 21, 22). However, it is the proportion of preventable PID in women with incident infection that is required to evaluate programs designed to lower infection rates.

Previous reviews (1, 23–25) of prospective data on progression to PID have been qualitative. Here we synthesize data from these studies to provide quantitative estimates of the difference in rates at which PID develops in women with and without CT. From these estimates, we derive first the probability that an episode of CT will cause an episode of PID and then the proportion of episodes of PID caused by CT, in women with CT, that could be prevented by randomly timed annual screening—that is, screening not associated with recent risk of infection, referred to herein as “annual screening.”

In this article, we first reappraise the results from prospective studies, including nonrandomized and uncontrolled studies. We then present a 3-state continuous-time Markov model, which allows for not only different rates of development of PID in CT-infected (CT+) and CT-uninfected (CT−) women but also spontaneous clearance of CT, a competing risk. It also allows for CT infection in women who are initially uninfected (CT−) and for reinfection with CT in women who are initially CT+ but who spontaneously clear infection or are tested and treated. We then reanalyze each of the controlled studies, both randomized and not, under the assumptions of the model, and finally synthesize the results altogether, to obtain estimates of the proportion of incident CT that results in PID and the proportion of CT-caused PID in women who become CT-infected that can be prevented by annual screening.

MATERIALS AND METHODS

Information sources and their interpretation

We identified 8 studies (Table 1) that followed women with CT to assess the proportion who developed PID in the absence of, or prior to, treatment (the search strategy is presented in Web Appendix 1, available at http://aje.oxfordjournals.org/). Some studies were set in sexually transmitted disease clinics, and others were based on infected women recruited by population screening. Five were observational, one of which, a clinic-based study, included a control group. The remainder were randomized controlled trials, all based on screened populations. In the Prevention of Pelvic Infection (POPI) Trial (11), participants in one arm were screened and treated immediately, while in the other arm screening and treatment were deferred until the end of the study.

Table 1 shows the proportion of women with PID in each arm at the end of each study. Most investigators have focused attention on these proportions and, in controlled studies, on relative risks. However, it is necessary to account for the mean follow-up time $t_o$ for each observation $o$. The last column of Table 1 shows crude study-specific estimates of the rates of progression to PID, based on a constant hazard within each study and arm.

The crude estimates of the progression rate in CT+ women vary by factors of 30-fold: Studies based on STI clinics have shorter follow-up periods and higher apparent progression rates, while those based on screened populations tend to have longer follow-up periods and lower progression rates. There are 3 possible explanations for this. Firstly, CT infections spontaneously resolve without treatment, so even if progression rates were constant over time, with CT clearance as a “competing risk,” the numbers of women acquiring PID must slowly decrease over time toward a background level seen in a representative control group. The effect of this will tend to be greater in studies with longer follow-up periods. However, the mean duration of untreated CT has been estimated to be over a year (17–20), so this would explain only a small part of the apparent differences. A second possible explanation is that women visiting STI clinics are more likely to have symptoms than those identified through screening, and the higher CT load associated with symptoms (26–28) may confer a greater risk of developing PID. A third possible explanation is that the risk of developing PID may be greater during the period immediately after CT infection (1, 29), even if the infection is asymptomatic.

There is a fundamental difference between STI clinic studies and screening studies. Women are most likely to visit STI clinics because of the onset of symptoms or because of concern following recent possible exposure to infection. They are therefore more likely to be recently infected. This is supported by evidence that symptoms mostly occur soon after infection (6, 20). Screening studies, by contrast, recruit participants with asymptomatic, prevalent CT infection who have already “survived” an unknown period of time without developing PID. Further, if the risk of PID is higher immediately following infection, the screening studies must underestimate the overall progression rate. For example, if asymptomatic infection lasted exactly 12 months and if 20% of incident cases developed PID (at which point they were treated for CT), always exactly 3 months after infection, the annual progression rate that would be observed in patients recruited from a prevalent population would be only 5.9% (i.e., 3/(0.8 × 12 + 0.2 × 3) × 0.2), a gross underestimate. In such a scenario, only screening and treatment occurring during the first 3 months of a 12-month infection would prevent PID. As a result, a single screen every year would prevent only 29.4% (5.9%/20% = 29.4%) of the PID cases caused by CT.

A model for progression to PID in the controlled prospective studies

In our model for prospective studies (Figure 1), women who are CT+ and CT− at the start of the study begin in
states 1 and 2, respectively. Women in state 1 may clear infection and progress to the CT− state, state 2, at rate $\lambda^C$, or they may develop PID (state 3) at a study-specific rate, $\theta^C+_{fts}$ for study s. Women in state 2 may become infected with CT, which occurs at rate $\lambda^I_s$, or may develop PID at rate $\theta^C/C_0s$ if they acquire or already have a non-CT infection which carries a PID risk. We assume that the prevalence and incidence rates for other STIs are the same in each arm. The difference $\delta$ between $\theta^C+_{fts}$ and $\theta^C/C_0s$ represents the rate of acquiring PID among women with a current CT infection that can be causally attributed to CT. The Figure 1 model is represented by the Markov transition rate matrix ($G$), shown in equation 1:

\[
G = \begin{bmatrix}
- (\lambda^C + \theta^{CT+}_s) & \lambda^C & \theta^{CT+}_s \\
\lambda^I_s & - (\lambda^I_s + \theta^{CT-}_s) & \theta^{CT-}_s \\
0 & 0 & 1
\end{bmatrix}.
\]  

Only comparative studies can contribute information on $\delta$, but we cannot derive estimates of $\delta$ from comparative studies without information on $\lambda^C$ and $\lambda^I_s$. We assume that the CT clearance rate for asymptomatic infection, $\lambda^C$, is constant across studies, at 0.74 (95% credible interval: 0.61, 0.89) per year, based on a synthesis of evidence from

### Table 1. Information From Prospective Studies That Followed Women With Chlamydia to Measure the Development of Pelvic Inflammatory Disease

<table>
<thead>
<tr>
<th>First Author, Year (Reference No.)</th>
<th>Study Design</th>
<th>Study Population</th>
<th>Arm</th>
<th>Data (Observation o), $r_o/n_o$</th>
<th>Crude Proportion, $p_o = r_o/n_o$</th>
<th>Follow-up Period, $t_o$</th>
<th>Crude Annual Rate of PID, $-(\log(1 - p_o))/t_o$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hook, 1994 (7)</td>
<td>Uncontrolled Clinic</td>
<td>CT+</td>
<td>3/93</td>
<td>0.032</td>
<td>0.012, 0.091</td>
<td>2 weeks</td>
<td>0.82, 0.30, 2.40</td>
</tr>
<tr>
<td>Geisler, 2008 (6)</td>
<td>Uncontrolled Clinic</td>
<td>CT+</td>
<td>2/115</td>
<td>0.017</td>
<td>0.005, 0.061</td>
<td>2 weeks</td>
<td>0.44, 0.14, 1.59</td>
</tr>
<tr>
<td>Paavonen, 1980 (8)</td>
<td>Uncontrolled Clinic</td>
<td>CT+</td>
<td>3/15</td>
<td>0.200</td>
<td>0.078, 0.481</td>
<td>1 month</td>
<td>2.79, 1.01, 8.21</td>
</tr>
<tr>
<td>Rahm, 1986 (9)</td>
<td>Uncontrolled Screened</td>
<td>CT+</td>
<td>4/102</td>
<td>0.048</td>
<td>0.016, 0.097</td>
<td>3 months</td>
<td>0.18, 0.07, 0.46</td>
</tr>
<tr>
<td>Rees, 1980 (10)</td>
<td>Controlled Clinic</td>
<td>CT+</td>
<td>8/67</td>
<td>0.119</td>
<td>0.063, 0.222</td>
<td>7–90 days</td>
<td>1.02, 0.52, 2.01</td>
</tr>
<tr>
<td>Oakeshott, 2010 (11) (POPI Trial)</td>
<td>RCT Screened Untreated CT+</td>
<td></td>
<td>7/74</td>
<td>0.093</td>
<td>0.047, 0.183</td>
<td>1 year</td>
<td>0.099, 0.049, 0.205</td>
</tr>
<tr>
<td>Scholes, 1996 (13)</td>
<td>RCT Screened</td>
<td>Unscreend</td>
<td></td>
<td>16/1,112</td>
<td>0.014</td>
<td>0.009, 0.023</td>
<td>0.014, 0.009, 0.024</td>
</tr>
<tr>
<td>Rahm, 1986 (9)</td>
<td>RCT Screened</td>
<td>Unscreend</td>
<td></td>
<td>1/63</td>
<td>0.016</td>
<td>0.004, 0.087</td>
<td>0.016, 0.004, 0.089</td>
</tr>
<tr>
<td>Rees, 1980 (10)</td>
<td>Controlled Clinic</td>
<td>CT−</td>
<td>14/1,128</td>
<td>0.011</td>
<td>0.007, 0.019</td>
<td>1 year</td>
<td>0.012, 0.007, 0.021</td>
</tr>
<tr>
<td>Oakeshott, 2010 (11) (POPI Trial)</td>
<td>RCT Screened</td>
<td>Unscreend</td>
<td></td>
<td>7/645</td>
<td>0.011</td>
<td>0.005, 0.022</td>
<td>0.011, 0.005, 0.022</td>
</tr>
<tr>
<td>Scholes, 1996 (13)</td>
<td>RCT Screened</td>
<td>Unscreend</td>
<td></td>
<td>44/645</td>
<td>0.068</td>
<td>0.051, 0.091</td>
<td>0.011, 0.005, 0.022</td>
</tr>
<tr>
<td>Østergaard, 2000 (12)</td>
<td>RCT Screened</td>
<td>Unscreend</td>
<td></td>
<td>20/487</td>
<td>0.041</td>
<td>0.027, 0.063</td>
<td>0.042, 0.027, 0.065</td>
</tr>
<tr>
<td>Scholes, 1996 (13)</td>
<td>RCT Screened</td>
<td>Unscreend</td>
<td></td>
<td>9/443</td>
<td>0.020</td>
<td>0.011, 0.039</td>
<td>0.021, 0.011, 0.039</td>
</tr>
<tr>
<td>Scholes, 1996 (13)</td>
<td>RCT Screened</td>
<td>Unscreend</td>
<td></td>
<td>43/867</td>
<td>0.050</td>
<td>0.037, 0.066</td>
<td>0.021, 0.011, 0.039</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; CT, Chlamydia trachomatis; CT+, CT-infected; CT−, CT-uninfected; PID, pelvic inflammatory disease; POPI, Prevention of Pelvic Infection; RCT, randomized controlled trial.
previous studies on CT duration (20). Similar estimates have been reported elsewhere (17). For CT incidence $\lambda$, we assume an infection rate of 5% per year in women who are CT− at the outset and a reinfection rate of 15% per year in women who are CT+ and tested and treated. Given an average duration of CT infection of 1–1.5 years, a 5% incidence reflects a CT prevalence of 5%–7.5%, which accords with the baseline CT prevalence in the 3 trials. Reinfection rates are higher by approximately a factor of 3 (30). We assume that all women in the Scholes et al. (13) and Østergaard et al. (12) (95% CT−) studies are at risk of infection, whereas all women in the Rees (10) (50% CT+) and POPI (11) (100% CT+) studies are assumed to be at risk of reinfection. Detailed sensitivity analyses (Web Appendix 2) show that conclusions are relatively insensitive to the values assumed.

Some of the screening studies recorded the proportion of women who were tested and treated for CT during follow-up, thus shortening their time at risk of CT-PID. This was 43% for the POPI Trial (11), 32% for Østergaard et al. (12), and 8.1% for Rees (10). The Scholes et al. paper (13) implies that some women were tested but does not report how many; we have assumed a proportion of 32%, following Østergaard et al. (12), but we explored the impact of a lower rate in sensitivity analyses (Web Appendix 2).

We consider 2 models. In the first, the causal rate of PID due to CT, δ, is constant throughout the duration of infection. In the second, the causal rate during the period immediately following CT infection (δ1) is different from the causal rate during the remainder of the infection (δ2). We fit models with the rate δ1 lasting for periods of 30, 60, and 90 days postinfection.

Estimation

Details of the estimation methods are given in Web Appendix 3. In brief, each study arm provided data with a binomial likelihood, which estimates a probability parameter, such as $p_{ij}(t)$, representing the study-specific transition probability that a person in state $i$ occupies state $j$, $t$ years later. These probabilities are entries in the transition probability matrix $P(t)$. If the transition rates are constant throughout the period $t$, then the transition probabilities, on which we have data (Table 1), are functionally related to the transition rates in equation 1 by Kolmogorov’s (31) forward equations:

$$dP_{ij}/dt = P(t) \times G. \tag{2}$$

Estimation is carried out using a Bayesian approach, where the posterior distribution is sampled through Markov chain Monte Carlo simulation. The rate parameters in equation 1 are given vague priors, the values of which are updated by the data. Numerical solutions to the forward equations are found in each Markov chain Monte Carlo cycle, using the WinBUGS (32) add-on WBDiff (33), which uses the Runge-Kutta method (34, 35).

From the POPI Trial (11), only the CT+ groups were included in this analysis, because the causal rate estimate is thereby based on a randomized comparison. The proportion of women with CT in the early treated arm are considered, in effect, to be CT− at the start of the observation period. Those in the deferred treatment group are, of course, CT+. However, in the Scholes et al. (13) and Østergaard et al. (12) trials, while the women in the screened arm are again taken to be CT− at the start of follow-up, those in the unscreened arm are a mixture of (untreated) CT+ and CT−. The study-specific CT prevalences in the screened arms are estimators of these mixing proportions.

Further details on how the observed data inform the rate parameters in the matrix $G$ (equation 1) are given in Web Appendix 3. The joint posterior distribution of the rate parameters is used to obtain posterior distributions for the 2 parameters of interest: the proportion of CT infections that result in PID and the proportion of CT-caused PID in women who become infected with CT that is preventable by annual screening.

Probability that incident CT will cause an episode of PID

While the comparative studies provide an estimate of the difference $\delta$ in progression rates between CT+ and CT− women, the target parameter of interest is the proportion of incident CT infections that result in PID. We assume that a proportion, $\phi$, of incident CT cases are symptomatic and treated, clearing at rate $\lambda$, and acquiring PID at the same rate as asymptomatic women. The remaining $1 - \phi$ infections are asymptomatic. To simplify the equations and maintain generality, we ignore non-CT-related PID as a competing risk at this stage. Even if the rate of non-CT PID were twice that observed in the screened arm of POPI, this would only change the value of our target parameter by about 2.5% (i.e., a multiplicative factor of 1.025). The probability that either type of infection will cause PID is the ratio of the rate of leaving state 1 for state 3 to the rate of leaving state 1, a standard result from competing-risks analysis (36). The proportion of incident cases of CT leading to an episode of PID in a homogenous model is therefore

$$\kappa = (1 - \phi) \times \frac{\delta}{\delta + \lambda} + \phi \times \frac{\delta}{\delta + \lambda^T}. \tag{3}$$

Calculation of $\kappa$ requires estimates of $\phi$ and $\lambda^T$. For the former, we use the estimate from Geisler et al. (6), 0.24 (95% credible interval: 0.17, 0.32). For the latter, we assume that treated infection lasts 4–8 weeks.

In the 2-stage piecewise homogenous model, with the initial rate lasting for $B$ days, the probability of developing PID in the first $B$ days and the probability of developing PID subsequently, conditional on having neither developed PID nor cleared infection in the first $B$ days, are summed:

$$\kappa = (1 - \phi) \times \left(1 - \exp\left(-\left(\lambda^T + \delta_1\right) \times \frac{B}{365}\right) \times \left(\frac{\delta_1}{\delta_1 + \lambda^T}\right) + \exp\left(-\left(\lambda^T + \delta_1\right) \times \frac{B}{365}\right) \times \left(\frac{\delta_1}{\delta_2 + \lambda^T}\right) + \phi \times \left(1 - \exp\left(-\left(\lambda^T + \delta_1\right) \times \frac{B}{365}\right) \times \left(\frac{\delta_1}{\delta_1 + \lambda^T}\right) + \exp\left(-\left(\lambda^T + \delta_1\right) \times \frac{B}{365}\right) \times \left(\frac{\delta_1}{\delta_2 + \lambda^T}\right)\right). \tag{4}$$
Estimation of the proportion of PID episodes prevented by screening

Even with 100% coverage, an annual screening program for CT would not prevent all cases of PID. In this paper, we estimate the proportion of episodes of CT-related PID that could be prevented in women who become infected with CT. The expressions for this quantity are derived and presented in Web Appendix 3. This applies to the benefits to CT+ women who are screened and treated and is not designed to measure the full effect of screening, since it does not take into account reduction in PID due to reduced CT incidence.

Plan of analysis

The data analysis is carried out in several steps in order to show the influence of the different types of data on the estimates of interest. Firstly, we examine each of the comparative studies separately. We then synthesize information from all 3 randomized controlled trials and then from the 4 comparative studies. Studies are combined on the basis that the progression rates in CT− women vary between studies in an arbitrary way, but the difference δ between the progression rates from CT− and CT+ has the same fixed effect across studies.

Estimation and model critique

Posterior mean values and 95% credible intervals were obtained from the Bayesian Markov chain Monte Carlo package WinBUGS, version 1.4.3 (32). Unless otherwise stated, vague priors were employed throughout, so that results are dominated by the data. Note that the prior structure for the rates is functionally equivalent to a model with a rate for PID caused by CT and a competing risk of non-CT-related PID. To assess goodness of fit, we used the posterior mean residual deviance, whose expected value is approximately equal to the number of data points under the assumption that the model is true (35, 37, 38). Further details are available in Web Appendix 4, and the WinBUGS code is provided in Web Appendix 5.

RESULTS

Table 2 shows the study-specific estimates from the homogenous Markov model of the difference δ, which we interpret as the causal rate of PID in CT+ women, and the proportion of incident CT infections that result in PID. While the POPI Trial data delivered the lowest estimates of both parameters, the credible intervals for the estimates from the other studies were considerably wider and approximately compatible with the POPI results. The estimates from the Scholes et al. study (13) were relatively insensitive to changes in assumptions on the proportion tested during the follow-up period (Web Appendix 2).

Table 3 presents the results of combining the information from the different studies under each of the homogenous and piecewise homogenous models. In all of the models, the mean residual deviance was close to the number of data points, indicating a good fit. In addition, it was apparent that the estimates under the 1-rate model were broadly similar for data 1) from the POPI Trial only, 2) pooled from all trials, and 3) including the observational study as well. The causal rate of PID ranged from 0.15 cases per year to 0.19 cases per year, and the probability of CT causing PID ranged from 0.12 to 0.16. The proportion of CT-caused PID that was preventable by annual screening remained at 61%-63%. The degree of overlap between these estimates and the shape of the posterior distributions can be seen in Figure 2.

The data did not distinguish the 1- and 2-rate models in terms of goodness of fit. The rate in the initial period is estimated to be approximately 50% higher than the rate during the remainder of infection, which is slightly lower than that for the 1-rate model. Both rates are insensitive to the assumed time period for the initial rate between 30 and 90 days (Table 3). Confidence intervals are very wide, however, and a Bayesian $P$ value testing the null hypothesis of equal rates in the 60-day model was only 0.67. The assumption of 2 rates has no real effect on the causal probability of PID. The estimated proportion of PID cases that can be prevented by screening was reduced marginally to 54%-58% for $B = 90$ days and $B = 30$ days, respectively, although the 95% credible intervals were wide (Figure 3).

Sensitivity analyses (Web Appendix 2) showed that although results from some of the individual studies can be sensitive to changes in assumed infection and reinfection rates, none of the results for the pooled analyses using the single-rate model were changed by more than a multiplicative factor of 4%. Results from the 2-rate model were more sensitive, with $\kappa$ varying between 0.15 and 0.20 when $B = 60$.

DISCUSSION

Early literature on the progression from CT to PID cited high rates based on uncontrolled studies in high-risk populations (8, 39). Over time, these estimates were increasingly questioned (4, 40, 41). Quantitative estimates of the causal

Table 2. Posterior Mean Values for the Causal Rate of Pelvic Inflammatory Disease (PID) in Chlamydia trachomatis (CT)-infected Women and the Probability That an Incident Case of CT Will Cause PIDa

| First Author, Year (Reference No.) | Causal Rate of PID in CT-infected Women, δ | Probability That CT Will Cause PID |
|-----------------------------------|------------------------------------------|
| Oakeshott, 2010 (11) (POPI Trial)  | 0.15, 0.02, 0.31 | 0.12, 0.02, 0.24 |
| Scholes, 1996 (13)                | 0.26, 0.03, 0.55 | 0.20, 0.03, 0.34 |
| Østergaard, 2000 (12)             | 1.03, 0.13, 2.49 | 0.43, 0.11, 0.65 |
| Rees, 1980 (10)                   | 0.76, 0.07, 1.78 | 0.37, 0.06, 0.59 |

Abbreviations: CrI, credible interval; CT, Chlamydia trachomatis; PID, pelvic inflammatory disease; POPI, Prevention of Pelvic Infection.

a Estimated using the Markov model shown in Figure 1.
The relation between CT and PID have been made using a wide range of study types (42), including linkage studies of treated women (14, 43–47), arguments from retrospective data (15), and the use of CT serology (48–51). More recently, it has been accepted that linkage studies cannot be used to generate reliable estimates (24). The approach presented here has some novel aspects. The Markov model characterizes the causal rate as the difference in progression rates in CT− and CT+ individuals. It takes account of the “competing” role of CT clearance and of the fact that persons testing CT+ upon screening have already “survived” a period without developing PID. Finally, we have allowed for the impact of CT treatment of women in the untreated arms of trials and likely reinfection rates among those who were treated.

The continuous-time Markov formulation avoids the limitations of traditional relative risk estimates (11–13) of the effect of CT screening. As the follow-up period is extended, the risk ratio estimate will be diluted by CT clearance and PID caused by infections (both CT and non-CT) acquired

Table 3. Estimates of the Proportion of Incident Chlamydia trachomatis (CT) That Results in Pelvic Inflammatory Disease (PID) and the Proportion of CT-caused PID Cases in CT-infected Women That Can Be Prevented by Annual Screening

<table>
<thead>
<tr>
<th>Method and Included Studies</th>
<th>Mean Residual Deviance</th>
<th>No. of Data Points</th>
<th>Causal Rate of PID</th>
<th>Probability That CT Will Cause Clinical PID</th>
<th>Proportion of Cases Prevented by Screening</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Posterior Mean</td>
<td>95% CrI</td>
<td>Posterior Mean</td>
</tr>
<tr>
<td>1-rate models</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>POPI data only (11)</td>
<td>2.0</td>
<td>2</td>
<td>0.15</td>
<td>0.02, 0.31</td>
<td>0.12</td>
</tr>
<tr>
<td>Trials only (11–13)</td>
<td>8.5</td>
<td>8</td>
<td>0.18</td>
<td>0.05, 0.33</td>
<td>0.15</td>
</tr>
<tr>
<td>All controlled studies (10–13)</td>
<td>10.6</td>
<td>10</td>
<td>0.19</td>
<td>0.06, 0.34</td>
<td>0.16</td>
</tr>
<tr>
<td>2-rate models</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All controlled studies—30 days</td>
<td>11.0</td>
<td>10</td>
<td>δ₁ = 0.28</td>
<td>0.02, 0.65</td>
<td>0.16</td>
</tr>
<tr>
<td>All controlled studies—60 days</td>
<td>10.9</td>
<td>10</td>
<td>δ₁ = 0.29</td>
<td>0.02, 0.65</td>
<td>0.16</td>
</tr>
<tr>
<td>All controlled studies—90 days</td>
<td>10.9</td>
<td>10</td>
<td>δ₁ = 0.28</td>
<td>0.02, 0.63</td>
<td>0.17</td>
</tr>
</tbody>
</table>

Abbreviations: CrI, credible interval; CT, Chlamydia trachomatis; PID, pelvic inflammatory disease; POPI, Prevention of Pelvic Infection.

Figure 2. Marginal posterior distributions of the probability that a Chlamydia trachomatis infection will cause clinical pelvic inflammatory disease (PID), estimated by each model. ACS, all controlled studies; POPI, Prevention of Pelvic Infection; RCT, randomized controlled trial.

after the screening and treatment, while in time the risk difference should converge to a stable estimate. Crude risk difference estimates (11, 22) should also be interpreted with caution. The POPI Trial data suggest that screening can prevent 83% of PID occurring during the following year in a prevalent CT+ population (11). Based on the Markov model, these data tell us that the proportion of incident CT-cased PID that can be prevented by annual screening is lower, between 55% and 65%.

The study has also shown for the first time that similar estimates can be derived, though indirectly and with less precision, from trials of screening versus no screening. The Scholes trial in particular has been criticized (21) because testing and treatment rates outside the protocol are unlikely to be the same in the 2 arms. We have used the information available, and sensitivity analyses, to mitigate this. Even with this defect, the trials are clearly superior sources of evidence on the causal effects of CT to the often-cited observational studies, where the difference in CT status at the outset is likely to correlate with risk factors that favor development of PID: higher partner change, greater risk of exposure to other agents that cause PID, and greater prior exposure to CT.

The study relates to “annual screening” at a random point in time. In practice, the benefit of screening for the individual infected woman may be greater if it occurs soon after partner change, as recommended in some countries, or less if partner testing and treatment are not pursued effectively.

The main weakness of this analysis lies in the reliance of almost all prospective studies on clinical PID as the outcome and between-study variation in case definition (42). While clinical PID reflects the cost to care providers of treating PID, it does not always signal the presence of the tubal inflammation (salpingitis) that is responsible for ectopic pregnancy and tubal infertility. The clinical definition may overestimate the rate of laparoscopically defined salpingitis, which may only be present in 45%–75% of cases (50, 52, 53). If the clinical PID that is not salpingitis is never caused by CT, then our estimate of the probability that CT causes PID can be interpreted as the probability that CT causes salpingitis. However, if the proportion of PID that is true salpingitis is the same regardless of whether it is caused by CT, then the probability that CT will cause an episode of salpingitis is likely to be lower than the estimates we derived, equaling about 40%–80% of our estimate.

Prospective studies may also underdiagnose PID, since case-control studies of tubal factor infertility suggest that up to 80% of women with tubal factor infertility report no history of PID (51, 54–56). However, Wölner-Hanssen (57) found that only 11% of women with tubal factor infertility reported never having had clinical symptoms. In prospective studies where participants are questioned about symptoms of PID, few cases are likely to be missed.

The systematic difference in observed PID rates in studies with shorter and longer follow-up periods has led to the suggestion that the progression rate is not constant over time (1). The potential impact of this has been explored in the mathematical modeling literature, using Markov or similar models (29, 58) to predict the proportion of PID prevented. However, with one exception (58), these studies have generated estimates of the proportion prevented from scenario analyses based on estimates of progression rates that were derived from models which do not allow for the competing risk of clearance or the left-truncation inherent in screening studies. The key contribution of this paper is to actually

![Figure 3. Marginal posterior distributions of the proportion of pelvic inflammatory disease (PID) episodes caused by *Chlamydia trachomatis* (CT) in women with CT that each model estimated would be prevented by annual screening. ACS, all controlled studies; POPI, Prevention of Pelvic Infection; RCT, randomized controlled trial.](https://academic.oup.com/aje/article-abstract/178/3/484/98355)
estimate all of these parameters consistently within the Markov model framework. We found that the proportion of CT-related PID that could be prevented by annual screening may be as low as 55%, which might represent only 20%–25% of all PID (59). Estimates based on 2-rate models were relatively insensitive to the duration of a period of heighten
ed PID risk following CT infection but were derived from a single, nonrandomized study. The sparse data available do not distinguish homogenous and piecewise homogenous models statistically, and the confidence intervals for the parameters are wide. However, both 1- and 2-rate models are considered plausible a priori (1).

REFERENCES


