Human Genome Epidemiology (HuGE) Review

Joint Effects of the N-Acetyltransferase 1 and 2 (NAT1 and NAT2) Genes and Smoking on Bladder Carcinogenesis: A Literature-based Systematic HuGE Review and Evidence Synthesis

Simon Sanderson1,2, Georgia Salanti2,3, and Julian Higgins2,3

1 Department of Public Health and Primary Care, Institute of Public Health, University of Cambridge, Cambridge, United Kingdom.
2 National Health Service Public Health Genetics Unit, Institute of Public Health, University of Cambridge, Cambridge, United Kingdom.
3 Medical Research Council Biostatistics Unit, Institute of Public Health, University of Cambridge, Cambridge, United Kingdom.

Received for publication April 25, 2006; accepted for publication March 30, 2007.

Bladder cancer is an increasingly important international public health problem, with over 330,000 new cases being diagnosed each year worldwide. In a systematic review and evidence synthesis, the authors investigated the joint effects of the N-acetyltransferase genes NAT1 and NAT2 and cigarette smoking on bladder carcinogenesis. Studies were identified through an exhaustive search of multiple electronic databases and reference lists and through direct contact with study authors and experts. Random-effects meta-analysis was used within a Bayesian framework to investigate individual effects of NAT1 and NAT2 acetylation status on bladder cancer risk, while a novel approach was used to investigate joint effects of these two genes with cigarette smoking. An increased risk of bladder cancer was found in NAT2 slow acetylators (odds ratio = 1.46, 95% credible interval (CI): 1.26, 1.68) but not in NAT1 fast acetylators (odds ratio = 1.01, 95% CI: 0.86, 1.22). The joint effects in the highest risk category (NAT2 slow acetylator, NAT1 fast acetylator, and current or ever cigarette smoking) as compared with the reference category (NAT2 fast acetylator, NAT1 slow acetylator, and never smoking) were associated with an odds ratio of 2.73 (95% CI: 1.70, 4.31). The importance of considering joint effects between genetic and environmental factors in the etiology of common complex diseases is underlined.

environmental exposure; genetics; genotype; meta-analysis; N-acetyltransferase 1; NAT2 protein, human; smoking; urinary bladder neoplasms

Abbreviations: CI, credible interval; DIC, deviance information criterion; GST, glutathione S-transferase; NAT, N-acetyltransferase.

Editor’s note: This article also appears on the website of the Human Genome Epidemiology Network (http://www.cdc.gov/genomics/hugenet/default.htm).

Bladder cancer

Bladder cancer is one of most common urologic malignancies worldwide, with approximately 330,000 new cases occurring each year. In the United Kingdom, bladder cancer is the fifth most common cancer, with 12,500 new cases and 5,000 deaths per year (1, 2). In the United States, it is the fourth commonest cancer in men and the ninth commonest in women, with 56,200 new cases and 12,600 deaths per year (3). Approximately 80 percent of bladder cancers occur in people aged 60 years or over, and as the population of the developed world ages, the incidence of bladder cancer will increase.

Correspondence to Dr. Simon Sanderson, Strangeways Research Laboratories, Institute of Public Health, University of Cambridge, Worts Causeway, Cambridge CB1 8RN, United Kingdom (e-mail: simon.sanderson@phgfoundation.org).
In developed countries, more than 90 percent of bladder cancers are urothelial transitional-cell carcinomas (1). Risk factors associated with bladder carcinogenesis include increasing age, male gender, and exposure to carcinogenic aromatic and heterocyclic amines (such as benzidine), either through occupation or through cigarette smoke (1, 3–5). Interest has focused on aromatic amines such as 4-aminobiphenyl, because they are found not only in cigarette smoke but also in several industrial compounds. Variants in several genes involved in the metabolism of these chemicals have also been investigated, including glutathione S-transferase M1 (GSTM1), N-acetyltransferase 1 (NAT1), and N-acetyltransferase 2 (NAT2) (6).

**NAT1 and NAT2 gene variants**

The NAT1 and NAT2 genes are both located on chromosome 8 (NAT1 on 8p21.3–23.1 and NAT2 on 8p21.3–23.1 and 8p22) (7, 8). An international committee standardized nomenclature for these genes in 1995, and known gene variants are listed and updated online (9, 10). NAT1*4 and NAT2*4 are the reference (wild-type) alleles. Over 25 NAT1 and NAT2 alleles have now been identified.

The NAT1 and NAT2 enzymes catalyze N-acetylation (usually deactivation) and O-acetylation (usually activation) of aromatic and heterocyclic amines and are particularly active in the liver, gastrointestinal tract, and urinary bladder, among other organs and tissues (7, 8, 11). The relative contributions of hepatic and extrahepatic activities of NAT1 and NAT2 in health and disease are not fully understood. Genetic variants of NAT1 and NAT2 are known to alter the metabolic rate of exogenous compounds such as caffeine and drugs such as isoniazid and dapsone. Classification of acetylation phenotype (slow or fast) can be made directly by measuring the metabolism of a probe drug or indirectly by measuring the genotype previously associated with that phenotype (7, 8). NAT1 and NAT2 variants are thought to modify human cancer risk by altering the rate at which potentially carcinogenic compounds are either neutralized or activated in different organs and tissues.

NAT1 and NAT2 variants are believed to have different effects on the risk of a number of human cancers, such as those of the bladder, colon, and breast. Slow NAT2 acetylation status is associated with an increased risk of bladder cancer, while fast NAT2 acetylation status is associated with colorectal cancer risk (7, 8, 12, 13). Although NAT1 has been associated with a number of cancers (including bladder cancer and colorectal cancer), findings are much less consistent than those for NAT2 (13, 14). A number of investigators have studied the joint effects of NAT1, NAT2, smoking, and occupation on bladder cancer risk, but published results have been inconsistent (14–17). These inconsistencies may be caused by differences in exposure to other genetic or environmental susceptibility factors interacting with NAT1, NAT2, and/or smoking, differences in methods of determining acetylation status, low statistical power, or selective reporting of positive findings.

Given this continuing uncertainty, we conducted a systematic, literature-based review of the individual effects of NAT1 and NAT2 and their joint effects with smoking on bladder carcinogenesis. We employed a novel approach to synthesizing evidence of joint effects, in which studies providing information on only one or two risk factors can still contribute relevant information on the three-way joint effects.

**METHODS**

This systematic review was conducted using the methods of the Human Genome Epidemiology Network (18) and the Cochrane Collaboration (19).

**Review inclusion and exclusion criteria**

Case-control and cohort studies investigating the association of bladder cancer with NAT1, NAT2, or both were eligible for the review. Studies without a disease-free control group were excluded.

**Search strategy and data extraction**

A comprehensive search of electronic databases was conducted, using a predetermined search strategy based on Medical Subject Headings (or equivalent thesaurus headings) and text words (see Appendix). Entrez PubMed, BIOSIS, EMBASE, and the Science Citation Index were searched from the earliest date of each database to December 2005. No language restrictions were imposed. Additional references were sought from published reviews, and the reference lists of identified reports were examined. Titles and abstracts of identified reports were scrutinized, and potentially relevant papers were retrieved to determine whether they met the inclusion criteria. Two reviewers (S. S. and G. S.) independently extracted data from the original articles using a preprinted proforma. A third reviewer (J. H.) was available for arbitration. The authors of primary studies were contacted in an attempt to identify unpublished studies, to obtain additional data, or to obtain clarification about study details.

Whenever cross-classification by smoking status was reported, this was recorded. We also recorded information about selection criteria for cases and controls; matching procedures; the method of collecting data on smoking history (interviewer or self-reported); blinded assessment of genotype and/or phenotype; conformity to Hardy-Weinberg equilibrium; explicit consideration of population stratification; demographic characteristics of participants, including ethnic origin, age, and gender; classification of acetylation status (based on genotype or phenotype); classification of smoking status; qualitative reporting of joint effects (for studies where the full joint effect was not studied); linkage disequilibrium between the NAT1 and NAT2 genes; and absolute numbers of bladder cancer cases and controls, cross-tabulated by NAT1 genotype status (slow or rapid), NAT2 genotype status (slow or rapid), and smoking status (yes or no), or any collapsed version of this $2 \times 2 \times 2 \times 2$ table that included both case/control status and one of the genes. NAT acetylation status was classified as either slow or rapid on the basis of direct phenotypic measurement or indirectly from the genotype, using the classification shown in table 1.
Smoking status was classified using dichotomizations reported in the primary studies.

Statistical methods

We first conducted random-effects meta-analyses of associations between NAT1 and bladder cancer and NAT2 and bladder cancer using the odds ratio as the effect measure. Pooled effect sizes were calculated using log odds ratios, and we defined the heterogeneity parameter as the standard deviation of the distribution of the study-specific log odds ratios. The following potential sources of heterogeneity were investigated: ethnicity (Caucasian and Indian vs. Asian), acetylation status classification method (phenotype vs. genotype), and source of the control population (hospital-based vs. population-based). A random-effects meta-regression model was fitted for each of these variables, assuming the same heterogeneity variance within each subgroup. The importance of each variable was assessed by expressing this heterogeneity variance as a proportion of the total heterogeneity across studies (an $R^2$ analog). The presence of possible publication bias was assessed graphically and statistically (20). It was not possible to investigate the impact of deviations from Hardy-Weinberg equilibrium, since genotype data were typically not available.

Using standard techniques, investigation of the joint effects of NAT1, NAT2, and smoking would be possible only in the very small subset of studies providing data on all exposures. Therefore, we estimated the joint effects of the three risk factors using a novel approach involving the whole data set (21). In brief, the goal of the analysis was to estimate simultaneously the odds ratios for each combination of NAT1, NAT2, and smoking status in comparison with a “low-risk” reference group of persons who were NAT1 rapid acetylators, NAT2 slow acetylators, and non-smokers. Studies not reporting the full $2 \times 2 \times 2$ table could contribute relevant information to the model if the data were supplemented with external information on the prevalence of the exposures not addressed by that study. We compared a model allowing freely varying joint effects with models 1) assuming no effect of NAT1 and 2) assuming multiplicative effects of the three exposures on the odds ratio scale (see below).

The novel synthesis method requires input in the form of prior distributions for prevalences of unobserved exposures. For smoking prevalence, we used country-specific data from the World Health Organization, giving each proportion a prior standard deviation of 1.5 percent to reflect uncertainty in the magnitude and relevance of the World Health Organization figures (22). For ethnicity-specific prevalences of NAT1 and NAT2, we used a combination of 1) other studies in the meta-analysis that allowed estimation of prevalence and 2) external studies of gene prevalence, identified by searching the database of the Human Genome Epidemiology Network (12). We assumed that the heterogeneity would be the same for each of seven log odds ratios.

We performed all analyses using Markov chain Monte Carlo methods in WinBUGS (23) within a Bayesian framework, taking advantage of the flexibility of WinBUGS as well as its ability to incorporate full uncertainty in all unknown parameters. To compare models, we calculated the deviance information criterion (DIC) for each model. A rule of thumb is that a difference of less than 2 in the DIC represents similarly supported models, with a DIC that is lower by 3 or more indicating a more appropriate model (23). Bayesian analyses yield credible intervals rather than confidence intervals; a 95 percent credible interval describes a range in which the probability that the unknown quantity lies within this interval, after seeing the data, is 95 percent. Approximately noninformative prior distributions are placed on the location parameters (such as the log odds ratios or the log prevalence). For the heterogeneity variance, we primarily used a half-normal prior with a variance of 1, and we further explored the impact of several alternative prior distributions (24). All analyses were based on a chain length of 50,000 after discarding the first 10,000 to allow for convergence, which was assessed by comparing chains with different initial values and observing the Brooks and Gelman plot (23). A full exposition of the above methods is available elsewhere (21).

RESULTS

After obtaining the full-text articles, we identified 36 studies meeting the inclusion criteria (see below), and we excluded six (25–30). We did not discover any unpublished studies, although we were informed of one study in progress. We received six replies from direct contacts with study authors.

Characteristics of the included studies

The 36 studies included (13–15, 31–63) contributed a total of 12,509 subjects for analysis (see table 2 or Web table 1 [www.aje.oxfordjournals.org] for full details). All were case-control studies. The studies were conducted in predominantly European Caucasian and Asian ethnic groups. Most of the cases were elderly and male, as would be expected from the epidemiology of bladder cancer. Control groups varied in their composition, ranging from healthy students to hospital inpatients. Earlier studies tended to use the phenotype to classify acetylation status, while later studies inferred the phenotype from the genotype. All included studies used histologic analysis to confirm bladder cancer. Most of the primary studies did not provide details on how

<table>
<thead>
<tr>
<th>Genotype(s)</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>At least one of the alleles NAT2<em>4, NAT2</em>12B-D, and NAT2*13, or a heterozygote of these alleles</td>
<td>Rapid acetylation All other slow acetylation genotypes</td>
</tr>
<tr>
<td>At least one of the alleles NAT1<em>10, NAT1</em>21, NAT1<em>24, and NAT1</em>25</td>
<td>Rapid acetylation All other slow acetylation genotypes</td>
</tr>
</tbody>
</table>
TABLE 2. Characteristics of studies included in a systematic review and meta-analysis of the joint effects of \(N\)-acetyltransferase 1 (\(NAT1\)), \(N\)-acetyltransferase 2 (\(NAT2\)), and cigarette smoking on bladder carcinogenesis

<table>
<thead>
<tr>
<th>First author (ref. no.) and year of publication</th>
<th>Total no. of subjects</th>
<th>Cases</th>
<th>Mean* (or median) age (years)</th>
<th>No. of males</th>
<th>% males</th>
<th>Controls</th>
<th>Mean (or median) age (years)</th>
<th>No. of males</th>
<th>% males</th>
<th>Notes</th>
<th>Gene(s) studied</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brockmoller (32), 1996</td>
<td>747</td>
<td>374</td>
<td>71</td>
<td>251</td>
<td>67</td>
<td>373</td>
<td>65.8</td>
<td>242</td>
<td>65</td>
<td>NAT2</td>
<td></td>
</tr>
<tr>
<td>Cartwright (33), 1982</td>
<td>318</td>
<td>111</td>
<td>Males: 68; females: 73</td>
<td>84</td>
<td>93.2</td>
<td>207</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cascorbi (13), 2001</td>
<td>768</td>
<td>425</td>
<td>Median 73</td>
<td>277</td>
<td>65</td>
<td>343</td>
<td>Median 65</td>
<td>220</td>
<td>64</td>
<td>NAT2<em>4; NAT2</em>10</td>
<td></td>
</tr>
<tr>
<td>Dewan (34), 1995</td>
<td>157</td>
<td>77</td>
<td>53.4</td>
<td>77</td>
<td>100</td>
<td>80</td>
<td>53.9</td>
<td>80</td>
<td>100</td>
<td>NAT2</td>
<td></td>
</tr>
<tr>
<td>Filiadis (35), 1999</td>
<td>236</td>
<td>89</td>
<td>65.8</td>
<td>75</td>
<td>84</td>
<td>147</td>
<td>63.9</td>
<td>122</td>
<td>83</td>
<td>NAT2</td>
<td></td>
</tr>
<tr>
<td>Garcia-Closas (36), 2005</td>
<td>2,299</td>
<td>1,150</td>
<td>66</td>
<td>1,004</td>
<td>65</td>
<td>1,002</td>
<td>65</td>
<td>1,002</td>
<td>87</td>
<td>NAT2*10; NAT2</td>
<td></td>
</tr>
<tr>
<td>Gu (37), 2005</td>
<td>1,020</td>
<td>507</td>
<td>63.8</td>
<td>395</td>
<td>80</td>
<td>513</td>
<td>62.7</td>
<td>395</td>
<td>77</td>
<td>NAT2*10; NAT2</td>
<td></td>
</tr>
<tr>
<td>Hanke (31), 1990</td>
<td>89</td>
<td>67</td>
<td></td>
<td></td>
<td></td>
<td>22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hanssen (38), 1985</td>
<td>147</td>
<td>105</td>
<td>72.3</td>
<td>81</td>
<td>77</td>
<td>42</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hao (39), 2004</td>
<td>157</td>
<td>69</td>
<td>58</td>
<td>56</td>
<td>81</td>
<td>88</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Horai (40), 1989</td>
<td>254</td>
<td>51</td>
<td>68.4</td>
<td>41</td>
<td>80</td>
<td>203</td>
<td>21</td>
<td>91</td>
<td>45</td>
<td>NAT2</td>
<td></td>
</tr>
<tr>
<td>Hsieh (15), 1999</td>
<td>258</td>
<td>74</td>
<td></td>
<td></td>
<td>78</td>
<td>184</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hung (41), 2004</td>
<td>415</td>
<td>201</td>
<td>NR</td>
<td>201</td>
<td>100</td>
<td>214</td>
<td>NR</td>
<td>214</td>
<td>100</td>
<td>NAT2*10; NAT2</td>
<td></td>
</tr>
<tr>
<td>Inatomi (42), 1999</td>
<td>231</td>
<td>85</td>
<td>66.3</td>
<td>70</td>
<td>82</td>
<td>146</td>
<td>63.3</td>
<td>84</td>
<td>58</td>
<td>NAT2</td>
<td></td>
</tr>
<tr>
<td>Ishizu (43), 1995</td>
<td>162</td>
<td>71</td>
<td>60.8</td>
<td>58</td>
<td>82</td>
<td>91</td>
<td>30.5</td>
<td>91</td>
<td>100</td>
<td>NAT2</td>
<td></td>
</tr>
<tr>
<td>Kaisary (45), 1987</td>
<td>205</td>
<td>95</td>
<td>69</td>
<td>68</td>
<td>72</td>
<td>110</td>
<td>64</td>
<td>79</td>
<td>72</td>
<td>NAT2</td>
<td></td>
</tr>
<tr>
<td>Karakaya (62), 1986</td>
<td>132</td>
<td>23</td>
<td>NR</td>
<td>17</td>
<td>74</td>
<td>109</td>
<td>NR</td>
<td>57</td>
<td>52</td>
<td>NAT2</td>
<td></td>
</tr>
<tr>
<td>Katoh (46), 1999</td>
<td>238</td>
<td>116</td>
<td>67.5</td>
<td>90</td>
<td>78</td>
<td>122</td>
<td>62.4</td>
<td>72</td>
<td>59</td>
<td>NAT2*10; NAT2</td>
<td></td>
</tr>
<tr>
<td>Kim (47), 2000</td>
<td>334</td>
<td>113</td>
<td>NR</td>
<td>93</td>
<td>82</td>
<td>221</td>
<td>NR</td>
<td>184</td>
<td>83</td>
<td>NAT2</td>
<td></td>
</tr>
<tr>
<td>Ladero (48), 1985</td>
<td>240</td>
<td>83</td>
<td>65.7</td>
<td></td>
<td></td>
<td>157</td>
<td>22.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower (49), 1979</td>
<td>378</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Miller (50), 1983</td>
<td>52</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mittal (63), 2004</td>
<td>211</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mommsen (51), 1985</td>
<td>328</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Okkels (52), 1987</td>
<td>496</td>
<td>254</td>
<td>69</td>
<td>133</td>
<td>52</td>
<td>242</td>
<td>64</td>
<td>118</td>
<td>49</td>
<td>NAT2*10; NAT2</td>
<td></td>
</tr>
<tr>
<td>Peluso (53), 1998</td>
<td>160</td>
<td>114</td>
<td></td>
<td></td>
<td>46</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Risch (54), 1995</td>
<td>248</td>
<td>189</td>
<td>NR</td>
<td>151</td>
<td>80</td>
<td>59</td>
<td>NR</td>
<td>48</td>
<td>81</td>
<td>No other details provided</td>
<td></td>
</tr>
<tr>
<td>Roots (55), 1989</td>
<td>202</td>
<td>101</td>
<td></td>
<td></td>
<td>101</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shnakenberg (56), 1998</td>
<td>214</td>
<td>60</td>
<td>72.5</td>
<td>42</td>
<td>70</td>
<td>154</td>
<td>37.6</td>
<td>89</td>
<td>58</td>
<td>No details provided</td>
<td></td>
</tr>
<tr>
<td>Su (57), 1998</td>
<td>87</td>
<td>27</td>
<td></td>
<td></td>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Taylor (14), 1998</td>
<td>433</td>
<td>230</td>
<td>65.3</td>
<td>178</td>
<td>77</td>
<td>203</td>
<td>63.3</td>
<td>167</td>
<td>82</td>
<td>No other details provided</td>
<td></td>
</tr>
<tr>
<td>Tsukino (58), 2004</td>
<td>650</td>
<td>325</td>
<td>69.4</td>
<td>255</td>
<td>78</td>
<td>325</td>
<td>67.1</td>
<td>255</td>
<td>78</td>
<td>NAT2</td>
<td></td>
</tr>
<tr>
<td>Vaziri (59), 2001</td>
<td>149</td>
<td>53</td>
<td>65.8</td>
<td>42</td>
<td>79</td>
<td>96</td>
<td>64.2</td>
<td>63</td>
<td>66</td>
<td>NAT2</td>
<td></td>
</tr>
<tr>
<td>Wang (60), 2002</td>
<td>51</td>
<td>17</td>
<td>66.4</td>
<td></td>
<td></td>
<td>34</td>
<td>67</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Woodhouse (61), 1982</td>
<td>67</td>
<td>30</td>
<td>70</td>
<td>20</td>
<td>67</td>
<td>27</td>
<td>77</td>
<td>13</td>
<td>48</td>
<td>Other information NR</td>
<td></td>
</tr>
<tr>
<td>Jaskula-Sztul (44), 2001</td>
<td>376</td>
<td>56</td>
<td>62.2</td>
<td>41</td>
<td>73</td>
<td>320</td>
<td>30.9</td>
<td>160</td>
<td>50</td>
<td>NAT2<em>3; NAT2</em>4; NAT2<em>10; NAT2</em>11; NAT2<em>14; NAT2</em>15</td>
<td></td>
</tr>
</tbody>
</table>

* Mean age is shown unless otherwise specified.
† NR, not reported.
smoking status was measured, the tools used, or how subjects were classified into subgroups and whether this was based on a standard definition.

In most of the studies, investigators did not report any attempt to control for confounding (population stratification), although a number of studies were conducted in apparently homogeneous populations. The majority of genotyping studies did not report whether or not gene frequencies were in Hardy-Weinberg equilibrium. Table 3 shows the distribution of data contributed from the included studies, depending on whether they had investigated NAT1, NAT2, or smoking alone or in combination. Only one study’s authors reported cross-classified data for NAT1, NAT2, and smoking status (14). Six studies included separate information for both genes or used two different study populations; therefore, these studies were included as two separate data sets (36, 37, 41, 49, 52, 60). Thus, in total, 42 data sets were available.

NAT1, NAT2, and bladder cancer risk: individual effects

The pooled odds ratio for NAT2 slow acetylation status and bladder cancer risk was 1.46 (95 percent credible interval (CI): 1.26, 1.68) (figure 1). The variance for the log odds ratio across studies (heterogeneity) was 0.31 (95 percent CI: 0.13, 0.53). Alternative different prior distributions for the heterogeneity parameter did not alter the estimated odds ratio or its credible interval, although the heterogeneity variance ranged from 0.24 to 0.35. Pooled analyses by ethnic group (Caucasian or Asian), control type (hospital- or population-based), and classification of acetylation status (genotype or phenotype) all showed consistent increases in risk of similar magnitudes and pointing in the same direction (figure 2).

For NAT1 rapid acetylation status, the corresponding odds ratio was 1.01 (95 percent CI: 0.86, 1.22) (figure 3). The variance of the log odds ratio was 0.14 (95 percent CI: 0.01, 0.41). The pooled effect size was not much affected by the different prior distributions; the largest increase was to an odds ratio of 1.07 (95 percent CI: 0.71, 1.56), while the heterogeneity point estimate varied substantially between 0.07 and 0.55, with similar changes in the credible intervals.

Analyses pooled by ethnic group and control source were similar, although there was a tendency towards higher odds ratios in Asian populations and in population controls (figure 2). However, these results were based on much smaller numbers of studies and total participants. For both NAT1 and NAT2, no substantial drop in heterogeneity was observed among the different subgroups.

No evidence of asymmetry in the funnel plot was observed for studies investigating NAT2; the regression coefficient was 0.01 (95 percent CI: −0.19, 0.22), as compared with a theoretical value of 0 in the absence of asymmetry. For NAT1, the small number of studies precluded a sound statistical assessment; however, there was no evidence of association between the effect size and the sample size, with a regression coefficient of −0.11 (95 percent CI: −0.38, 0.11).

NAT1, NAT2, and cigarette smoking: joint effects

Estimates of the prevalence of each gene, required for the novel joint-effects model and derived from a Human Genome Epidemiology review of NAT genes in colorectal cancer (12), are presented in table 4. The odds ratios for the full joint-effects model and their 95 percent credible intervals are presented in table 5.

The odds ratio for the highest risk category (current smoker, NAT1 rapid acetylator, and NAT2 slow acetylator) was 2.73 (95 percent CI: 1.70, 4.31), providing a quantification of the joint effects of these three risk factors on bladder carcinogenesis. The DIC for this model (a measure for comparing the suitability of different models) was 1,107; DICs were 1,120 for a multiplicative interaction model and 1,111 for a model assuming that NAT1 plays no role. This provides strong evidence for a departure from a no-multiplicative-interaction model and some evidence for a role of NAT1 in the context of the effects of NAT2 and smoking.

Heterogeneity for the log odds ratios in the main model was 0.57 (95 percent CI: 0.49, 0.77). We performed sensitivity analyses to investigate the impact of different degrees of information on the prevalence priors (noninformative prior, informative prior as assessed in table 1, and informative prior increasing 10 times the precision) and for the heterogeneity of the prevalence (noninformative prior and informative prior as estimated in table 1). Extreme values for key variables were also explored (such as a smoking

### Table 3. Distribution of exposure data contributed by the primary studies to the joint-effects model in a systematic review and meta-analysis of the joint effects of N-acetyltransferase 1 (NAT1), N-acetyltransferase 2 (NAT2), and cigarette smoking on bladder carcinogenesis

<table>
<thead>
<tr>
<th>NAT1</th>
<th>NAT2</th>
<th>Smoking</th>
<th>No. of studies</th>
<th>First author and ref. no</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>1</td>
<td>Taylor (14)</td>
</tr>
<tr>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>4</td>
<td>Cascorbi (13), Hsieh (15), Katoh (46), Jaskula-Stulz (44)</td>
</tr>
<tr>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>13</td>
<td>Wang (60), Hung (41), Mittal (63), Inatomi (42), Brockmoller (32), Risch (54), Dewan (34), Ishizu (43), Gu (37), Karakaya (62), Tsukino (58), Garcia-Closas (36), Roots (55)</td>
</tr>
<tr>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>2</td>
<td>Gu (37), Hung (41)</td>
</tr>
<tr>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>19</td>
<td>Okkels (52), Hao (39), Lower (49), Peluso (53), Cartwright (33), Mommsen (51), Miller (50), Kaisary (45), Horai (40), Hanssen (38), Schnakenberg (56), Hanke (31), Su (57), Woodhouse (61), Ladero (48), Filladis (35), Vaziri (59)</td>
</tr>
<tr>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>3</td>
<td>Okkels (52), Wang (60), Garcia-Closas (36)</td>
</tr>
</tbody>
</table>

* Lower et al. (49) reported results for two separate populations (Sweden and Denmark) in the same paper; these populations were included in the analysis separately.
† Patients were matched on smoking history in this study; thus, direct information about joint effects in the model could not be provided.
FIGURE 1. Relation between \( N \) acetyltransferase 2 (\( \text{NAT2} \)) acetylation status and bladder cancer risk in a systematic review and meta-analysis. "Lower D" represents the Danish portion of the study by Lower et al. (49); "Lower S" represents the Swedish portion. Horizontal lines, 95% credible interval.

FIGURE 2. Relation between \( N \) acetyltransferase 1 (\( \text{NAT1} \)) acetylation status and bladder cancer risk in a systematic review and meta-analysis. Horizontal lines, 95% credible interval.
prevalence of 90 percent). No substantial changes in the results of the analysis were observed.

DISCUSSION

This systematic review provides the most recent and comprehensive assessment of the effects of NAT1, NAT2, and smoking in bladder carcinogenesis. We developed a Bayesian approach to modeling joint effects between genes and other environmental exposures, based on analysis of the published literature (as opposed to individual patient data). This study confirmed previously reported evidence of an increased risk of bladder cancer in patients with NAT2 slow acetylation status. This increased risk was consistent between patients of Caucasian and Asian origin and was not influenced by methodological considerations, such as whether hospital or population controls were used or how acetylation status was determined. There was little support for fast NAT1 acetylation status alone in increasing bladder cancer risk, although this was based on a much smaller number of studies than those for NAT2.

However, NAT1 fast acetylation status may be important when NAT2 slow acetylation status and cigarette smoking are analyzed as joint effects; the odds ratio was 2.73 (95 percent CI: 1.7, 4.3) for people with adverse versions of all three exposures. This is an important result, because it emphasizes the importance of considering the contribution of

### TABLE 4. Results from a meta-analyses of the prevalence of the N-acetyltransferase 1 (NAT1) and N-acetyltransferase 2 (NAT2) rapid acetylation genotypes from external studies

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Caucasians and Indians (5 studies)</th>
<th>Asians (3 studies)</th>
<th>Caucasians and Indians (9 studies)</th>
<th>Asians (13 studies)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Prevalence</td>
<td>Heterogeneity (r)</td>
<td>Prevalence</td>
<td>Heterogeneity (r)</td>
</tr>
<tr>
<td>NAT1</td>
<td>Posterior mean 95% CI</td>
<td>Posterior mean 95% CI</td>
<td>Posterior mean 95% CI</td>
<td></td>
</tr>
<tr>
<td>Prevalence</td>
<td>0.48</td>
<td>0.74</td>
<td>0.43</td>
<td>0.40</td>
</tr>
<tr>
<td>Heterogeneity (r)</td>
<td>0.30, 0.67</td>
<td>0.27, 1.88</td>
<td>0.39, 0.47</td>
<td>0.24, 0.57</td>
</tr>
</tbody>
</table>

* Studies excluded from the present review and meta-analysis; that is, those studies obtained from the paper by Brockton et al. (12).
† CI, credible interval.
‡ The heterogeneity refers to the logit(prevalence).
TABLE 5. Odds ratio for bladder cancer according to category of cigarette smoking and N-acetyltransferase 2 (NAT2) and N-acetyltransferase 1 (NAT1) acetylation status in a random-effects synthesis of data from all studies*

<table>
<thead>
<tr>
<th>NAT2 acetylation status</th>
<th>NAT1 acetylation status</th>
<th>Odds ratio</th>
<th>95% credible interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>No current regular smoking</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rapid</td>
<td>Slow</td>
<td>1†</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rapid</td>
<td>0.83</td>
<td>0.36, 1.75</td>
</tr>
<tr>
<td></td>
<td>Slow</td>
<td>0.98</td>
<td>0.52, 1.62</td>
</tr>
<tr>
<td></td>
<td>Rapid</td>
<td>1.12</td>
<td>0.52, 1.98</td>
</tr>
<tr>
<td>Current regular smoking</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rapid</td>
<td>Slow</td>
<td>1.71</td>
<td>1.01, 2.83</td>
</tr>
<tr>
<td></td>
<td>Rapid</td>
<td>1.36</td>
<td>0.81, 2.14</td>
</tr>
<tr>
<td>Slow</td>
<td>Slow</td>
<td>2.36</td>
<td>1.47, 3.71</td>
</tr>
<tr>
<td></td>
<td>Rapid</td>
<td>2.73</td>
<td>1.70, 4.31</td>
</tr>
</tbody>
</table>

† Reference category.

multiple genetic and environmental factors in the etiology of complex diseases, especially when the results of single gene-association studies appear to be “negative.” This conclusion assumes that these exposures occur independently, which we believe to be justifiable, as we are not aware of evidence associating NAT genotype with smoking, although there is some evidence of partial linkage disequilibrium between the NAT1*10 allele and NAT2 rapid acetylation haplotypes (13).

Evidence that NAT2 slow acetylation does not confer the same risk in nonsmokers as it does in smokers provides evidence of a likely causal effect of smoking through a Mendelian randomization argument (64), since NAT2 status is unlikely to be associated with confounders. Within the Bayesian analysis, we can derive estimates of the main effect of the NAT2 slow acetylation genotype in smokers and nonsmokers, assuming a NAT1 prevalence of 50 percent. These odds ratios are 1.21 (95 percent CI: 0.93, 1.70) and 1.67 (95 percent CI: 1.23, 2.17), respectively, providing support for the notion of causality.

In an attempt to mitigate the problem of publication bias, we used a thorough search strategy, supplemented by direct contact with study authors, and used graphical and statistical methods of assessment. Although the response from authors was low, it should be borne in mind that a number of these studies were conducted before 1980, and the authors may have moved on, retired, or died. The small number of NAT1 studies makes formal assessment of publication bias unreliable, although there was no evidence of such an effect. Another potential source of bias is reporting bias in published studies, where authors selectively report results on the basis of their perceived interest and relevance; this is much more difficult to assess from the published literature.

Assessment of each primary study’s susceptibility to bias (often referred to as “quality”) and its impact on the results of systematic reviews is a difficult issue. We have used a number of key indicators to assess this, derived from recent consensus statements about gene-disease associations and gene prevalence and from observational epidemiologic studies. Although the main aim of these statements is study reporting, useful lessons about potential biases in study design, conduct, and execution can also be learned. However, it is not clear whether or not today’s standards are applicable to studies published over 35 years ago, given our increased understanding of genetics, epidemiology, statistics, and critical appraisal. In line with best practice, we did not use summary, quantitative scoring systems for rating studies. We assessed each study on its own merits and in context and presented the data so that readers can make their own assessments of the primary studies.

Three particular issues merit further attention: misclassification, population stratification, and Hardy-Weinberg equilibrium.

Misclassification of key study variables (exposures, outcomes, and confounders) is a potential problem in any epidemiologic study, but its effects can be compounded in a meta-analysis. All of the studies we included used histologic analysis to confirm bladder cancer, so the risk of outcome misclassification is likely to have been very low. However, there are potential problems in misclassification of three key variables: acetylation status, smoking history, and genotype.

Assessment of phenotype using probe drugs has a long track record and is considered to be reasonably valid and reliable, especially for NAT2 when caffeine is used as the substrate. The direct assessment of NAT1 status is more difficult (7, 8). Acetylation phenotype is affected by a number of other factors such as diet, concurrent administration of other drugs, and comorbid conditions, as well as other key liver enzyme systems (such as the cytochrome P-450 complex) which also influence the metabolism of xenobiotics used as probes for classification. As a result, acetylation phenotypes exhibit continuous and overlapping variability and thus cannot always be easily reduced to distinct categories, such as fast and slow. Blinding of assessors to a subject’s case/control status is an important consideration, because this knowledge may influence the conduct of the test.

Genotype has the advantage of being fixed at conception, and it can usually be determined with a high degree of validity and reliability. A small subset of NAT2 alleles (designated WT, M1–M4) can lead to misclassification of the inferred phenotype because of technical limitations in their determination (7, 8). Other potential genotyping pitfalls include the need to use specific polymerase chain reaction primers for both NAT1 and NAT2 and correctly determining phase. These problems mean that large sample sizes must be used, especially when investigating joint effects. As with the phenotypic measurement, blinding of assessors to a subject’s case/control status is also an important consideration.

The other key exposure is smoking history. Most of the primary studies we included did not provide details on how smoking status was measured, the tools used, or how subjects were classified into subgroups. Interviewers will almost certainly have not been blind to the case/control status of interviewed subjects, and self-report questionnaires may not be answered truthfully; both approaches could be subject to recall bias. We attempted to mitigate this problem by using a “current/ever” versus “never” classification for the definition of smoking status. However, we were unable
to determine whether there was also a dose-response relation for heavy smokers versus light smokers.

Most of the authors of review studies did not report any attempt to control for population stratification. For gene-disease associations, this source of confounding has been a relatively recent topic of interest, so most investigators would not necessarily have considered this (although it is interesting that one of the earliest studies matched subjects on nationality (49)). While a number of studies were conducted in apparently homogenous populations, this was not reported as an explicit strategy for controlling confounding. By using a broad classification of ethnic origin, we found a consistent effect size and direction for \textit{NAT2} slow acetylators, despite gene frequencies for \textit{NAT2} that varied widely between different ethnic groups. On the basis of this classification, an admixture of Caucasians and Asians (but not Caucasians and Indians) might be expected to lead to confounding; however, the results remained stable in the sensitivity analysis.

In the majority of genotyping studies included, investigators did not report whether or not their observed gene frequencies were in Hardy-Weinberg equilibrium. Violation of Hardy-Weinberg equilibrium is an important issue, because it may indicate underlying genotyping error in primary studies and may introduce bias into the results of genetic association studies and subsequent meta-analyses (65).

In this study, we have demonstrated that it is possible to investigate and model joint effects between genetic and environmental factors from the published literature using Bayesian methods. Evidence for a contribution of \textit{NAT2} slow acetylation status alone to bladder carcinogenesis is supported, while \textit{NAT1} rapid acetylation status is not. However, there is some evidence for a joint effect of these risk factors in bladder carcinogenesis.

Acknowledgments

The authors thank Dr. Iain Tatt, who conducted the initial planning and literature searches, and Dr. Ian White for statistical advice.

Conflict of interest: none declared.

References


APPENDIX
Search Strategy (Entrez PubMed)
1. N-Acetyltransferase
2. NAT1
3. NAT2
4. Arylamine N-acetyltransferase
5. Arylamine N-acetyltransferase [MeSH*]
6. Aromatic amine
7. Heterocyclic amine
8. Smoking
9. Smoking [MeSH]
10. Tobacco
11. Tobacco [MeSH]
12. Environment
13. Exposure
14. Environmental exposure [MeSH]
15. Bladder
16. Bladder [MeSH]
17. Urinary
18. Urothelial
19. Cancer
20. Neoplasm
21. Neoplasms [MeSH]
22. Carcino
23. Carcinoma [MeSH]
24. Tumour
25. Tumor
26. Tumour [MeSH]
27. DNA adduct
28. DNA adducts [MeSH]
29. Bladder neoplasms [MeSH]
30. Urologic neoplasms [MeSH]

*MeSH, Medical Subject Headings.