Research

Sarah Damery, Linda Nichols, Roger Holder, Ronan Ryan, Sue Wilson, Sally Warmington, Helen Stokes-Lampard and Kaveh Manavi

Assessing the predictive value of HIV indicator conditions in general practice:

a case–control study using the THIN database

Abstract

Background
UK HIV guidelines identify 37 clinical indicator conditions for adult HIV infection that should prompt an HIV test. However, few data currently exist to show their predictive value in identifying undiagnosed HIV.

Aim
To identify symptoms and clinical diagnoses associated with HIV infection and assess their relative importance in identifying HIV cases, using data from The Health Improvement Network (THIN) general practice database.

Design and setting
A case–control study in primary care.

Method
Cases (HIV-positive patients) were matched to controls (not known to have HIV). Data from 939 cases and 2576 controls were included (n = 3515). Statistical analysis assessed the incidence of the 37 clinical conditions in cases and controls, and their predictive value in indicating HIV infection, and derived odds ratios (ORs) for each indicator condition.

Results
Twelve indicator conditions were significantly associated with HIV infection; 74.2% of HIV cases (n = 697) presented with none of the HIV indicator conditions prior to diagnosis. The conditions most strongly associated with HIV infection were bacterial pneumonia (OR = 47.7; 95% confidence interval [CI] = 5.6 to 404.2) and oral candidiasis (OR = 29.4; 95% CI = 6.9 to 125.5). The signs and symptoms most associated with HIV were weight loss (OR = 13.4; 95% CI = 5.0 to 36.0), pyrexia of unknown origin (OR = 7.2; 95% CI = 2.8 to 18.7), and diarrhoea (one or two consultations).

Conclusion
This is the first study to quantify the predictive value of clinical diagnoses related to HIV infection in primary care. In identifying the conditions most strongly associated with HIV, this study could aid GPs in offering targeted HIV testing to those at highest risk.

Keywords
diagnosis; general practice; guidelines; HIV; primary care; database; THIN.

INTRODUCTION

An estimated 91 500 people in the UK were living with HIV at the end of 2010.[1] Treatments such as highly active anti-retroviral therapy (HAART) have significantly reduced the number of HIV-related deaths, and people living with diagnosed HIV can expect a near-normal life expectancy.[2] A key prognostic factor is the stage at which HIV is diagnosed,[3] with earlier diagnosis associated with improved patient survival,[4–6] reduced direct care costs,[7] and lower rates of onward transmission.[8,9] It is estimated that approximately 24% of all adult HIV infections in the UK are still undiagnosed.[1]

Recent efforts have been directed towards introducing universal HIV testing for asymptomatic individuals considered at high risk of infection.[10] UK national guidelines recommend routine HIV testing for general medical admissions and newly registering general practice patients in areas where HIV prevalence exceeds 2 per 1000 among 15- to 59-year-olds.[1] However, many patients may present symptomatically.[11] Evidence suggests that up to one-third of newly diagnosed patients may have had multiple encounters with primary care services,[12] with symptoms that, in hindsight, were clearly related to HIV.[13,14] Yet the clinical diagnosis of HIV-related conditions in primary care is complex, as many of the presenting signs or symptoms associated with HIV are commonly seen in people without it.

The UK HIV guidelines identify 37 unique clinical indicator conditions for adult HIV infection (Box 1), whose presence should prompt the offer of an HIV test.[10] While each of these indicator diagnoses has been reported in HIV-infected patients, few data currently exist to show their predictive value in identifying undiagnosed HIV. A predictive tool, based on patient presentation in primary care with any of the 37 clinical indicator conditions, has the potential to improve the detection of the symptoms most predictive of HIV infection. This could facilitate the implementation of targeted HIV testing in primary care settings where HIV testing is not routinely offered but where patients presenting with one or more indicator conditions are regularly encountered.

This study used anonymised patient records from The Health Improvement Network (THIN) general practice database, to identify the symptoms and clinical diagnoses associated with increased risk of HIV infection, in order to develop a predictive model to identify the relative importance of these factors in identifying HIV cases.

S Damery, PhD, research fellow; S Wilson, PhD, professor of clinical epidemiology; S Warmington, senior trial manager, Primary Care Clinical Sciences, University of Birmingham. L Nichols, PhD, statistician; R Holder, PhD, head of statistics; H Stokes-Lampard, PhD, FRCGP, clinical senior lecturer, GP, deputy director Primary Care Clinical Research and Trials Unit, Primary Care Clinical Sciences, University of Birmingham. R Ryan, MSc, research fellow, Public Health, Epidemiology and Biostatistics, University of Birmingham, Birmingham. K Manavi, MD, consultant physician, lead of HIV service, University Hospitals Birmingham NHS Foundation Trust, Birmingham.

How this fits in

Up to one-third of patients newly diagnosed with HIV may have had multiple encounters with primary care services, with symptoms that were, in hindsight, clearly related to HIV infection. National HIV guidelines identify 37 unique clinical indicator conditions whose presence should prompt the offer of an HIV test, but few data currently exist to show the predictive value of these conditions in indicating undiagnosed HIV. This study found the conditions most strongly associated with HIV diagnosis were bacterial pneumonia, oral candidiasis, and herpes zoster, suggesting that patients presenting in primary care with these conditions should be routinely offered an HIV test. However, only a minority of HIV patients show any indicator conditions prior to diagnosis, and most are asymptomatic. Nevertheless, many missed opportunities for earlier diagnosis may exist where patients present symptomatically, thus great scope exists for reducing late diagnosis by improving the testing practices of GPs.

METHOD

Data source

The THIN general practice database (http://www.thin-uk.com) contains individual longitudinal electronic patient records from 386 UK general practices (6.9 million patients). Participating practices regularly upload anonymised patient-level data to THIN. The database includes the demographic details of current and past patients and all medical diagnoses, treatments, prescriptions, clinical measurements, and laboratory results initiated by the practice, as well as referrals to hospital specialists and hospital diagnoses recorded by the general practices. Details of specific general practices and all patient identifiers are fully anonymised by THIN before the release of datasets for research purposes.

Model inclusion criteria

The THIN model was developed using a nested cohort study design, in which cases (patients known to be diagnosed with HIV or AIDS) were matched to controls (not known to have HIV or AIDS). All patients in the THIN database, aged 18 years and over, who were diagnosed with HIV or AIDS at least 1 year after registering with their practice were eligible for inclusion as cases in the study cohort. These were matched by 5-year age group and sex to controls (case to control ratio of 1:3 where possible) who were registered at the same practice during the year the case received their diagnosis, but who were not known to be HIV positive.

Variables included in the THIN model

Clinical Read codes for the conditions

Box 1. HIV clinical indicator conditions[a]

Respiratory medicine
- Bacterial pneumonia
- Aspergillosis

Neurology
- Aseptic meningitis/encephalitis
- Cerebral abscess
- Space-occupying lesion of unknown cause
- Guillain–Barré syndrome
- Transverse myelitis
- Peripheral neuropathy
- Dementia
- Leukoencephalopathy

Dermatology
- Severe or recalcitrant seborrhoeic dermatitis
- Severe or recalcitrant psoriasis
- Multidermatomal or recurrent herpes zoster

Gastroenterology
- Oral candidiasis
- Oral hairy leukoplakia
- Chronic diarrhoea of unknown cause
- Weight loss of unknown cause
- Salmonella, shigella or campylobacter
- Hepatitis B infection
- Hepatitis C infection

Oncology
- Anal cancer or anal intraepithelial dysplasia
- Lung cancer
- Seminoma
- Head and neck cancer
- Hodgkin's lymphoma
- Castleman's disease

Gynaecology
- Vaginal intraepithelial neoplasia
- Cervical intraepithelial neoplasia grade 2+

Haematology
- Any unexplained blood dyscrasia (including thrombocytopenia, neutropenia, lymphopenia)

Ophthalmology
- Infective retinal diseases including herpes viruses and toxoplasma
- Any unexplained retinopathy

Ear, nose, and throat
- Chronic parotitis
- Lymphoepithelial parotid cysts

Other
- Mononucleosis-like syndrome
- Pyrexia of unknown origin
- Any lymphadenopathy of unknown cause
- Any sexually transmitted infection

[a]The national guidelines[10] list 38 clinical indicators for possible HIV infection. However, 'lymphadenopathy of unknown cause' appears twice, so was treated as a single condition.

Table 1. Operational definition of specific HIV indicator conditions

Condition in national guidelines[10]	Operational definition used in the THIN model
Space-occupying lesion of unknown cause	Space-occupying lesion
Peripheral neuropathy	Excluded Read codes attributing condition to diabetes
Dementia	Dementia diagnosed under the age of 65 years
Severe or recalcitrant seborrhoeic dermatitis	Number of consultations for seborrhoeic dermatitis in specified period
Severe or recalcitrant psoriasis	Number of consultations for psoriasis dermatitis in specified period
Multidermatomal or recurrent herpes zoster	Herpes zoster
Chronic diarrhoea of unknown cause	Number of consultations for diarrhoea in specified period
Weight loss of unknown cause	Weight loss
Any unexplained blood dyscrasia	Blood dyscrasia
Any unexplained retinopathy	Retinopathy in a non-diabetic patient
Lymphadenopathy of unknown cause	Lymphadenopathy
Chronic parotitis	Parotitis

Figure 1. Exclusions made to THIN data.

included in the model were derived from the 37 unique indicator conditions described in the national guidelines.[10] Conditions with highly specific definitions that were unlikely to be coded as such in general practice were redefined (Table 1). In addition to the clinical indicator conditions, sociodemographic data (Townsend deprivation quintile, sex, age, ethnic distribution of each patient's area of residence, and an urban/rural residence indicator variable), and data regarding whether a patient had consulted their GP about more than one of the indicator conditions within a 12-month period, were obtained.

Analysis of the THIN model

The final statistical model was developed by entering the full set of clinical variables into a backward stepwise conditional logistic regression model, with the significance level for removal of indicator conditions set at 0.2. Model outputs were the characteristics of HIV cases and controls with regard to the frequency with which each HIV indicator condition was observed in each group, and odds ratios (ORs) with 95% confidence intervals (CIs) for each indicator condition. Modelling was undertaken using Stata software (version 11).
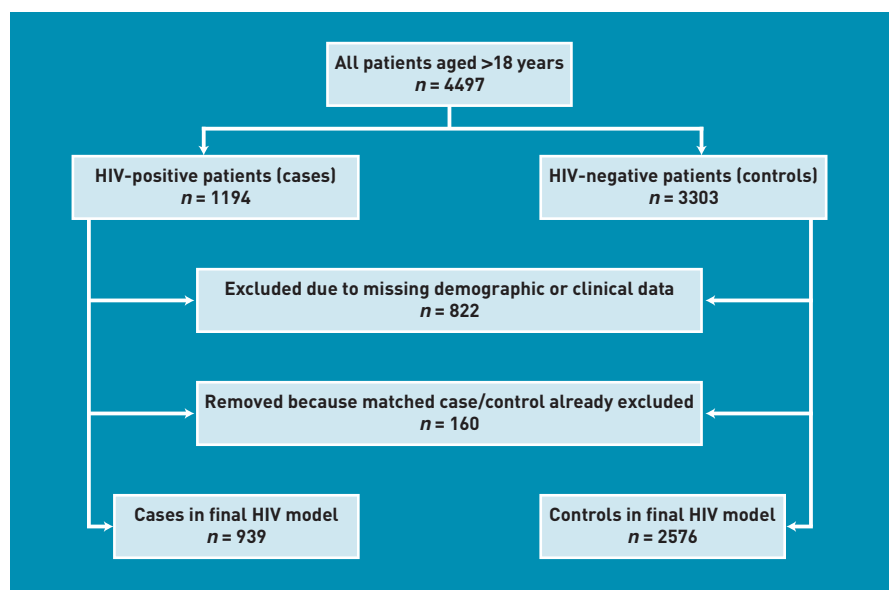
RESULTS

Clinical and patient characteristics

A total of 4497 patients aged 18 years or over (1194 HIV cases and 3303 matched controls), from 362 UK general practices, were eligible for inclusion in the THIN model. The date of HIV diagnosis for cases ranged from January 1989 to September 2010. After exclusions, 939 cases and 2576 controls remained in the final model (Figure 1). Although the initial goal was to match three controls to each case, this was not possible for every single case; thus the case to control ratio achieved in the final dataset was 1:2.74. The clinical characteristics of cases and controls are described in Table 2.

The majority of HIV cases were male (68.7%), with a mean age at diagnosis of 41 years (standard deviation [SD] = 11.4 years). The most common conditions diagnosed in the year before HIV diagnosis were diarrhoea, with 67 cases having at least one episode (7.1%), oral candidiasis ($n = 35$ cases; 3.7%), herpes zoster (shingles) ($n = 32$ cases; 3.4%), lymphadenopathy ($n = 28$ cases; 3.0%), and weight loss ($n = 27$ cases; 2.9%). The majority of the HIV cases ($n = 697$; 74.2%) had none of the HIV indicator conditions, and 20.8% had just one of the conditions ($n = 195$). For 16 of the HIV indicator conditions, there were no instances recorded in the year before HIV was diagnosed.

Sociodemographic characteristics

Table 3 outlines the sociodemographic characteristics of cases and controls. HIV cases were significantly more likely to live in the most deprived Townsend quintile in comparison to control patients ($n = 246$ cases; 26.2% versus $n = 534$ controls; 20.7%: $P<0.001$). The profile of case and control patients with regard to the proportion of residents in the same census ward who had been registered as Asian or Asian British in the 2001 census was similar across all quintiles ($P = 0.95$). For both

Table 2. Clinical characteristics of HIV cases and matched controls (THIN model)

	Cases		Controls	
	n	%	*n*	%
Total *n*	939		2576	
Males	645	68.7	1771	68.8
Mean age (standard deviation)	41	(11.4)	42	(11.0)
BHIVA conditions[a]				
Bacterial pneumonia	17	1.8	1	0.0
Aseptic meningitis/encephalitis	2	0.2	0	0.0
Peripheral neuropathy	2	0.2	1	0.0
Seborrhoeic dermatitis				
1 consultation only	8	0.9	6	0.2
2 consultations only	0	0.0	0	0.0
3+ consultations	0	0.0	0	0.0
Psoriasis				
1 consultation only	8	0.9	9	0.4
2 consultations only	7	0.8	4	0.2
3+ consultations	2	0.2	1	0.0
Herpes zoster	32	3.4	4	0.2
Oral candidiasis	35	3.7	2	0.1
Oral hairy leukoplakia	3	0.3	0	0.0
Diarrhoea				
1 consultation only	50	5.3	38	1.5
2 consultations only	11	1.2	6	0.2
3+ consultations	6	0.6	0	0.0
Weight loss	27	2.9	6	0.2
Salmonella, shigella or campylobacter	3	0.3	4	0.2
Hepatitis B infection	4	0.4	0	0.0
Hepatitis C infection	2	0.2	0	0.0
Anal cancer/anal intraepithelial dysplasia	1	0.1	0	0.0
Seminoma	0	0.0	1	0.0
Non-Hodgkin's lymphoma	4	0.4	1	0.0
Blood dyscrasia	9	1.0	3	0.1
Retinopathy in non-diabetic patient	1	0.1	0	0.0
Lymphadenopathy	28	3.0	7	0.3
Parotitis	2	0.2	1	0.0
Pyrexia of unknown origin	17	1.8	7	0.3
Any sexually transmitted infection	15	1.6	3	0.1
Number of indicator conditions per patient				
None of the indicator conditions	697	74.2	2477	96.2
One of the indicator conditions only	195	20.8	93	3.6
Two or more indicator conditions	47	5.0	6	0.2

[a]There were no cases of aspergillosis, cerebral abscess, space-occupying lesion, Guillain–Barré syndrome, transverse myelitis, dementia, leukoencephalopathy, lung/head and neck cancer, Castleman's disease, vaginal intraepithelial neoplasia, cervical intraepithelial neoplasia ≥grade 2, Infective retinal diseases, lymphoepithelial parotid cysts, or mononucleosis-like syndrome, in either case or control groups. BHIVA = British HIV Association.

cases and controls, around half resided in areas with the highest proportion of Asian or Asian British individuals (*n* = 466 cases; 49.6% versus *n* = 1282 controls; 49.8%). Similarly, more than half of both cases and controls were patients who lived in areas with the highest proportion of Black or Black British residents (*n* = 493 cases; 52.5% versus *n* = 1332 controls; 51.7%), and the profile of cases and controls across quintiles did not differ statistically (*P* = 0.78).

Results of the predictive model

All HIV indicator conditions, and some

Table 3. Sociodemographic variables for HIV cases and controls

	Cases		Controls	
	n	%	*n*	%
Townsend deprivation index				
(least deprived) quintile 1	135	14.4	544	21.1
quintile 2	145	15.4	426	16.5
quintile 3	182	19.4	533	20.7
quintile 4	231	24.6	539	20.9
(most deprived) quintile 5	246	26.2	534	20.7
Proportion of ward population that was Asian or Asian British at last census				
(lowest proportion) quintile 1	64	6.8	161	6.3
quintile 2	86	9.2	218	8.5
quintile 3	113	12.0	332	12.9
quintile 4	210	22.4	583	22.6
(highest proportion) quintile 5	466	49.6	1282	49.8
Proportion of ward population that was Black or Black British at last census				
(lowest proportion) quintile 1	37	3.9	108	4.2
quintile 2	111	11.8	317	12.3
quintile 3	90	9.6	292	11.3
quintile 4	208	22.2	527	20.5
(highest proportion) quintile 5	493	52.5	1332	51.7

Comparison of distribution of cases and controls: Townsend deprivation index P<0.001, proportion of ward population that was Asian or Asian British P = 0.95, proportion of ward population that was Black or Black British P = 0.78.

Table 4. Predictive value of HIV clinical indicator conditions in the THIN model[a]

Indicator condition	Odds ratio	Standard error	z	*P*-value	95% CI	
Statistically significant conditions						
Bacterial pneumonia	47.7	52.0	3.54	<0.001	5.6	404.2
Oral candidiasis	29.4	21.8	4.57	<0.001	6.9	125.5
Herpes zoster	25.4	14.2	5.76	<0.001	8.4	76.1
Weight loss	13.4	6.7	5.15	<0.001	5.0	36.0
Non-Hodgkin's lymphoma	12.6	15.0	2.13	0.033	1.2	129.8
Lymphadenopathy	11.3	5.3	5.15	<0.001	4.5	28.3
Sexually transmitted infection	10.8	7.6	3.38	0.001	2.7	43.2
Pyrexia of unknown origin	7.2	3.5	4.05	<0.001	2.8	18.7
Blood dyscrasia	5.7	4.0	2.44	0.015	1.4	22.9
Diarrhoea — one consultation only	3.7	0.9	5.48	<0.001	2.3	6.0
Diarrhoea — two consultations	4.4	2.3	2.81	0.005	1.6	12.1
Conditions that were not statistically significant[b]						
Parotitis	8.6	11.0	1.68	0.093	0.7	106.1
Psoriasis — one consultation only	2.6	1.5	1.69	0.091	0.9	7.9
Psoriasis — two consultations	3.0	2.5	1.38	0.168	0.6	14.8

[a]Also included in conditional logistic regression: Townsend deprivation index and proportion of ward that was Asian and/or Asian British (quintiles). [b]These conditions remained in the stepwise regression model, as their significance was below the 0.2 cut-off threshold for inclusion. Other HIV indicator conditions not shown in the table were excluded during the stepwise process.

sociodemographic data, were entered into the stepwise conditional logistic regression model (Table 4), in order to determine the factors that were most strongly associated with being an HIV case. Twelve conditions were retained in the final model: 10 of these had statistically significant ORs. The

condition most strongly associated with HIV diagnosis was bacterial pneumonia (OR = 47.7; 95% CI = 5.6 to 404.2): patients coded as having bacterial pneumonia were almost 50 times more likely to be diagnosed with HIV in the subsequent 12 months than patients without this condition. Patients with oral candidiasis were almost 30 times more likely to be diagnosed with HIV (OR = 29.4; 95% CI = 6.9 to 125.5), followed by those with herpes zoster (OR = 25.4; 95% CI = 8.4 to 76.1). The clinical signs and symptoms associated with subsequent HIV diagnosis were weight loss (OR = 13.4; 95% CI = 5.0 to 36.0), pyrexia of unknown origin (OR = 7.2; 95% CI = 2.8 to 18.7), and diarrhoea (OR = 3.7 for one consultation only; OR = 4.4 for two consultations).

Model sensitivity and specificity
The probability of being an HIV case was determined for each individual, using the clinical and sociodemographic model parameters in order to establish the sensitivity and specificity of the predictive model. From these data, the sensitivity and specificity of the model were established. An assessment of the model's predictive ability over a range of threshold values is shown in the receiver operating characteristic (ROC) curve in Figure 2.

The area under the curve was 0.66, with an optimum cut-off point (the point at which the combined model sensitivity and specificity were maximised) of 0.26. At this threshold, the model would correctly identify 42% of HIV cases, and correctly determine around 80% of controls as not having HIV.

DISCUSSION
Summary
This is the first study to develop a model that uses patient presentation in primary care with any of the indicator conditions from the current national guidelines for HIV testing, in order to quantify the predictive values of clinical diagnoses that may be related to HIV infection in the primary care setting. The THIN model included data on 939 HIV cases and 2576 matched controls, and identified 12 clinical diagnoses significantly associated with HIV infection.

The most common HIV indicator conditions recorded in the year before HIV diagnosis were diarrhoea (at least one consultation), oral candidiasis, herpes zoster, lymphadenopathy, and weight loss. Notwithstanding this, 74.2% of HIV cases had presented with none of the HIV indicator conditions, and only 20.8% of cases had just one of the conditions. There were no
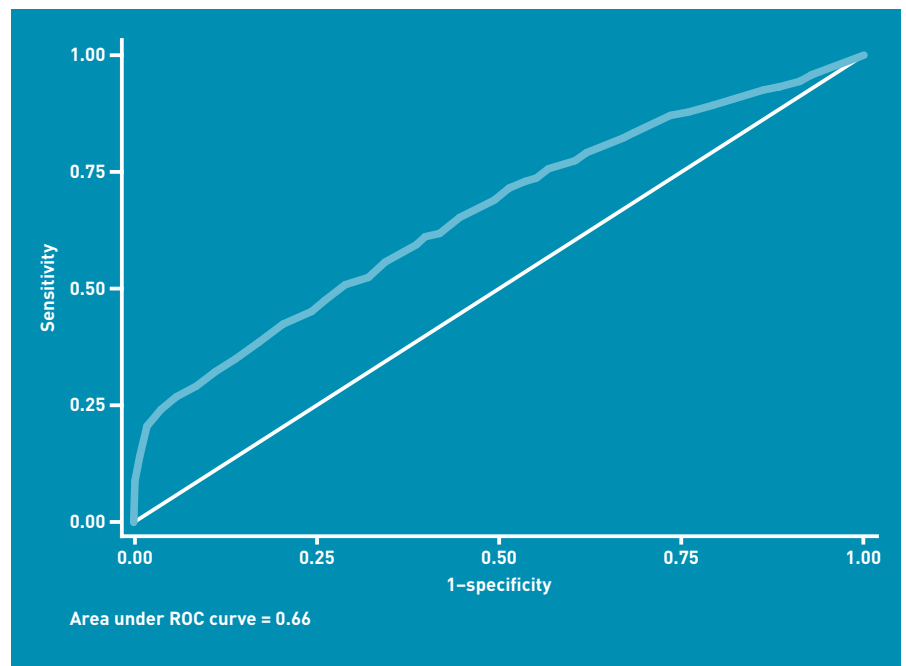
Figure 2. An ROC curve for the model developed using the THIN dataset.

diagnoses recorded in the year before HIV was diagnosed for 16 of the conditions listed in the guidelines. The conditions most strongly associated with subsequent HIV diagnosis were bacterial pneumonia, oral candidiasis, and herpes zoster. The signs and symptoms most strongly associated with subsequent HIV were weight loss, pyrexia of unknown origin, and diarrhoea (one or two consultations). The THIN model suggested an optimum cut-off threshold of 0.26 (sensitivity 42%; specificity 80%).

Strengths and limitations

This study has demonstrated the feasibility of using routinely collected general practice data to model the predictive capacity of the clinical conditions in the national HIV guidelines, for indicating potentially undiagnosed HIV infection. The model was developed using a large cohort of cases and controls, and case to control matching allowed a robust assessment of the utility of each of the HIV indicator conditions in predicting HIV infection. The majority of the HIV indicator conditions are seen relatively infrequently in general practice, so may not be useful for GPs in determining whether HIV testing should be offered to their patients. In identifying the guideline indicator conditions that are most strongly associated with HIV infection, this study could aid GPs in offering targeted HIV testing to those at highest risk, as indicator conditions with larger ORs must be considered for HIV testing in any setting.

This study also highlights areas of the national HIV guidelines where further research is needed. Diagnoses that were not represented in the THIN data, or that had a small OR, need to be investigated for the rate of HIV seropositivity in patients with these relatively rare conditions.

The THIN database is subject to the limitations inherent in all routinely collected data, which may include inconsistencies, variation between GPs, and missing data. There may be selection bias due to general practices electing to participate in THIN voluntarily, and there may be under-recording of symptoms and HIV diagnoses. Similarly, the selection of a limited number of clinical Read codes that corresponded to each of the 37 HIV indicator conditions may have reduced the ascertainment of these indicators. However, there is evidence that the reporting behaviour of general practices in THIN approximates to the national level.[15] All patients within a general practice contributing to THIN are included in the database, which minimises the possibility of patient-level selection bias. The THIN database is representative of the UK population,[16] and the representativeness of practices participating in THIN has also been demonstrated.[17] Only if under-recording of symptoms was particularly prevalent in the case or control group would bias arise in this respect, but there is no reason to believe that this was the case in the present study. Thus if there was under-recording of symptoms, these would

be expected to occur proportionately in both case and control groups, which would not alter the observed OR for each HIV indicator condition, but which may reduce the statistical significance of associations.

Any study using GP consultations as a means of quantifying individual patient risk is subject to the assumptions that, first, symptomatic patients are actually registered with a GP, and secondly, that symptomatic patients did indeed visit their GP to discuss their symptoms. However, a high proportion of patients may not consult their GP, even if showing symptoms of a potentially serious condition. Similarly, this study assumed that patients with HIV diagnoses were always known to the practice. This may not be the case if patients diagnosed with HIV outside of primary care (for example in a genitourinary medicine clinic) are unknown to the GP. If this is the case, the model developed for this study may underestimate the true prevalence of the HIV indicator conditions in the general population, and may consequently underestimate their predictive significance in indicating potentially undiagnosed HIV.

Comparisons with existing literature
As far as the authors are aware, this is the first published study of its kind in this field.

Implications for research and practice
Future research should validate the model in the general practice setting, with a large number of practices, so that validation could be undertaken with a high degree of reliability. This process should include the requirement for general practices to record the diagnostic outcome for all patients that the predictive model identifies as potentially carrying undiagnosed HIV, and who consequently undergo HIV testing. This would allow the sensitivity and specificity of the predictive model to be validated in the 'real-life' setting, and in locations with different levels of underlying HIV prevalence. Consideration of the financial (cost-effectiveness) implications of HIV testing was outside the scope of this study. Future research should undertake a full cost-effectiveness analysis to assess the utility of targeted HIV screening of symptomatic individuals within the primary care setting.

Given that one-third of HIV infected individuals are undiagnosed, and one-third of those newly diagnosed present late in the course of their infection, there is a need to expand HIV testing in primary care. Many missed opportunities for earlier diagnosis may exist, where patients present symptomatically; thus great scope exists for reducing late diagnosis by improving the testing practices of GPs, especially when they see patients with chronic diarrhoea, unexplained weight loss, or other symptoms that could be associated with seroconversion to HIV. The findings of this study could lead to a more accurate and potentially more cost-effective strategy for diagnosis of HIV-infected patients in primary care.

REFERENCES

1. Health Protection Agency. *HIV in the United Kingdom: 2011 report.* London: HPS, Colindale, 2011.

2. May M, Gompels M, Delpech V, *et al.* Impact of late diagnosis and treatment on life expectancy in people with HIV-1: UK Collaborative HIV Cohort (UK CHIC) Study. *BMJ* 2011; **343:** d6016.

3. Chadborn TR, Baster K, Deppech VA, *et al.* No time to wait: how many HIV infected homosexual men are diagnosed late and consequently die? *AIDS* 2005; **19(5):** 513–520.

4. Stohr W, Dunn DT, Porter K, *et al*, on behalf of the UK CHIC study. CD4 cell count and initiation of antiretroviral therapy: trends in seven UK centres, 1997–2003. *HIV Med* 2007; **8(3):** 135–141.

5. Krentz HB, Auld MC, Gill MJ. The high cost of medical care for patients who present late with HIV infection. *HIV Med* 2004; **5(2):** 93–98.

6. Brown AE, Kall M, Smith RD, *et al.* Auditing national HIV guidelines and policies: the United Kingdom CD4 surveillance scheme. *Open AIDS J* 2012; **6:** 149–155.

7. Marks G, Crepaz N, Janssen RS. Estimating sexual transmission of HIV from persons aware and unaware that they are infected with the virus in the USA. *AIDS* 2006; **20(10):** 1447–1450.

8. Granich RM, Gilks CF, Dye C, *et al.* Universal voluntary HIV testing with immediate antiretroviral therapy as a strategy for elimination of HIV transmission: a mathematical model. *Lancet* 2009; **373(9657):** 48–57.

9. Attia S, Egger M, Muller M, *et al.* Sexual transmission of HIV according to viral load and antiretroviral therapy: systematic review and meta-analysis. *AIDS* 2009; **23(11):** 1397–1404.

10. British HIV Association, British Association of Sexual Health and HIV, British Infection Society. *UK national guidelines for HIV testing 2008.* London: BHIVA, 2008.

11. Palfreeman A, Fisher M, Ong E, on behalf of the HIV Testing Guidelines Writing Committee. Testing for HIV: concise guidance. *Clin Med* 2009; **9(5):** 471–476.

12. Sudarshi D, Pao D, Murphy G, *et al.* Missed opportunities for diagnosing primary HIV infection. *Sex Transm Infect* 2008; **84(1):** 14–16.

13. Sullivan AK, Curtis H, Sabin CA. Newly diagnosed HIV infections: review in the UK and Ireland. *BMJ* 2005; **330(7503):** 1301–1302.

14. Donaldson L, Beasley C. *Improving the detection and diagnosis of HIV in non-HIV specialities including primary care.* London: Department of Health, 2007.

15. Blak BT, Thompson M, Dattani H, Bourke A. Generalisability of The Health Improvement Network (THIN) database: demographics, chronic disease prevalence and mortality rates. *Inform Prim Care* 2011; 19(4): 251–255.

16. European Observatory on Health Care Systems. *Health care systems in transition, United Kingdom.* Copenhagen: European Observatory on Health Care Systems, WHO Regional Office for Europe, Denmark, 1999.

17. Zakai NA, Katz R, Hirsch C, *et al.* A prospective study of anemia status, haemoglobin concentration and mortality in an elderly cohort: the cardiovascular health study. *Arch Intern Med* 2005; **165(19):** 2214–2220.