Can an electronic prescribing system detect doctors who are more likely to make a serious prescribing error?

Jamie J Coleman[1,2] • Karla Hemming[1] • Peter G Nightingale[2] • Ian R Clark[2] • Mary Dixon-Woods[3] • Robin E Ferner[1] • Richard J Lilford[1,2]

[1]College of Medical and Dental Sciences, University of Birmingham, Birmingham, UK
[2]University Hospitals Birmingham NHS Foundation Trust, Edgbaston, Birmingham, UK
[3]Department of Health Sciences, 2nd Floor Adrian Building, University of Leicester, Leicester, UK
Correspondence to: Jamie J Coleman. Email: Jamie.coleman@uhb.nhs.uk

DECLARATIONS

Competing interests
All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) (JJC). JJC, IRC, and PGN work within the University Hospital Birmingham NHS Foundation Trust which is collaborating with CSE Healthcare Systems to commercialise the PICS system in the UK. All other authors report no financial relationships with commercial entities that might have an interest in the submitted work; no

Summary

Objectives We aimed to assess whether routine data produced by an electronic prescribing system might be useful in identifying doctors at higher risk of making a serious prescribing error.

Design Retrospective analysis of prescribing by junior doctors over 12 months using an electronic prescribing information and communication system. The system issues a graded series of prescribing alerts (low-level, intermediate, and high-level), and warnings and prompts to respond to abnormal test results. These may be overridden or heeded, except for high-level prescribing alerts, which are indicative of a potentially serious error and impose a 'hard stop'.

Setting A large teaching hospital.

Participants All junior doctors in the study setting.

Main outcome measures Rates of prescribing alerts and laboratory warnings and doctors' responses.

Results Altogether 848,678 completed prescriptions issued by 381 doctors (median 1538 prescriptions per doctor, interquartile range [IQR] 328–3275) were analysed. We identified 895,029 low-level alerts (median 1033 per 1000 prescriptions per doctor, IQR 903–1205) with a median of 34% (IQR 31–39%) heeded; 172,434 intermediate alerts (median 196 per 1000 prescriptions per doctor, IQR 159–266), with a median of 23% (IQR 16–30%) heeded; and 11,940 high-level 'hard stop' alerts. Doctors vary greatly in the extent to which they trigger and respond to alerts of different types. The rate of high-level alerts showed weak correlation with the rate of intermediate prescribing alerts (correlation coefficient, $r = 0.40$, $P = <0.001$); very weak correlation with low-level alerts ($r = 0.12$, $P = 0.019$); and showed weak (and sometimes negative) correlation with propensity to heed test-related

warnings or alarms. The degree of correlation between generation of intermediate and high-level alerts is insufficient to identify doctors at high risk of making serious errors.

Conclusions Routine data from an electronic prescribing system should not be used to identify doctors who are at risk of making serious errors. Careful evaluation of the kinds of quality assurance questions for which routine data are suitable will be increasingly valuable.

spouses, partners, or children of the authors have relationships with commercial entities that might have an interest in the submitted work; none of the authors have non-financial interests that may be relevant to the submitted work.

Funding
The National Institute for Health Research (NIHR) Collaborations for Leadership in Applied Health Research and Care (CLAHRC) for Birmingham and Black Country and the Department of Health Policy Research Programme supported the undertaking of this work. The views expressed in this work do not necessarily reflect those of the funders

Ethical approval
Not applicable

Guarantor
JJC

Contributorship
All authors had full access to statistical reports and tables in the study and take responsibility for the

Introduction

The use of routine data for monitoring quality in health systems is well established. The advantages are many: the data are readily available and can be used at far less cost than prospectively designed studies.[1] Much of the academic literature has focused on use of routine data for making comparisons across hospitals and enabling detection of variation in process measures and outcomes.[2] The increasing use of large-scale information technology (IT) systems across healthcare presents new opportunities to use routine data to direct quality monitoring and improvement activities *within* organizations.

The absence of data on individual professional practice has been identified as a significant barrier to greater physician involvement in quality improvement[3] yet the potential of using routine data for this purpose has remained largely unevaluated. An important question concerns whether some doctors are more likely to make serious errors than others.[4] One possibility is that individuals who demonstrate a pattern of multiple low-level or moderate-level errors are at increased risk of making a higher-level, more consequential, error. If individuals at higher risk of making dangerous errors could be identified, they could be offered additional monitoring and support. Electronic systems that capture routine data about practitioners' behaviours provide opportunities for exploring whether different types of behaviours may be correlated with different outcomes.

Prescribing practice is a particularly good area in which to focus such study both because it is an important source of preventable harm[5–7] and because prescribing errors with a low risk of resulting in harm are much more frequent, at a population level, than serious errors.[8,9] We aimed to identify the extent to which routine prescribing data might be useful in identifying individuals who are at higher risk of making a serious prescribing error.

Methods

Setting and study population

The study was carried out in a large NHS Foundation Trust with two teaching hospital sites. The Trust has a locally-developed electronic prescribing system known as PICS (Prescribing, Information and Communication System), which is in use throughout all (approximately 1200) inpatient beds and for all prescribing except some chemotherapy regimens.

The system was first installed in the renal unit more than a decade ago,[10] and now covers general and specialist medical and surgical specialties apart from obstetrics, paediatrics and mental health. A key feature of the system, for purposes of our study, is that it provides decision support by generating messages alerting prescribers to potential problems. Programmed into the system are approximately 5000 alerts for contraindications, 3400 alerts relating to dose limits, and 1800 for drug interactions. The algorithms that generate the rules are based on the British National Formulary (BNF),[11] but are locally configured and updated regularly by a committee of medical and pharmacy specialists.

Computer-generated alerts relating to *prescriptions* are graded as follows (examples in Table 1):

- Low-level alerts, where the user is requested to 'tick a box', indicating that the message has been considered;
- Intermediate alerts, where the user must supply a password before the prescription can be continued;

accuracy of the data analysis. All authors contributed to the writing of the manuscript, the interpretation of data, and approved the final version. KH, PGN and IRC had access to the raw data

Acknowledgements

The authors are grateful for the support of the Birmingham Clinical Research Academy. The authors thank Dave Thompson and Ian Young for assistance with data extraction from the audit database. The authors are also grateful to James Reason and Frank Davidoff for their independent reviews and discussion of the contents of the manuscript

Table 1
Examples of types of prescribing alerts with descriptions and hierarchy of alerts

Hierarchy of alerts		*Types of alerts*		
		Drug contraindication	*Drug-drug interaction*	*Dose-range checking*
Low-level (minor severity)	Description	Used when the presence of a coded contraindication (usually disease state) means that the use of the medicine could be harmful	Used for combinations of medicines which in certain circumstances could cause harm	Use of a drug probably outside clinically acceptable or usual limits for dose quantity or timing, but risk of harm not considered high
	Specific system example	Atorvastatin (a cholesterol-lowering drug) should be used caution with renal impairment due to the risk of side-effects of muscle inflammation	The use of atenolol and diuretics (both drugs used in hypertension) together may cause low blood pressure	Daily dose warning limit for beclometasone nose drops is 18 drops
Intermediate (moderate severity)	Description	Used when the presence of a coded contraindication (usually disease state) means that the use of the medicine may be potentially hazardous and requires avoidance or other precautions	Used for potentially hazardous interactions of drugs which require avoidance or other special cautions	Use of a drug outside clinically acceptable limits for dose quantity or timing based upon the likelihood of severe harm with that dose the usual dosing of that drug
	Specific system example	Aspirin should not be given in the presence of gastrointestinal ulceration due to the risk of bleeding	Using methotrexate and ciclosporin (both immunosuppresants) together can lead to toxicity	The infusion rate of vancomycin (antibiotic) is limited to 600 mg/h; a higher rate may cause 'red man' syndrome or hypotension
High-level messages (hard stop)	Description	Used when the presence of a coded contraindication (usually disease state) prohibits the prescription due to risk of harm	Used for combinations of drugs that should be avoided in all situations	Use of a drug significantly outside clinically acceptable limits for dose quantity or timing based upon the likelihood of serious harm with that dose
	Specific system example	Abciximab (a potent drug used to prevent thrombosis in cardiac patients) is not prescribable to patients with brain tumours due to high risk of bleeding	The antibiotic erythromycin and the antipsychotic pimozide are disallowed in combination due to risk of potentially fatal cardiac rhythm abnormality	Amphotericin (Abelcetan) – an antifungal drug – > 6 mg/kg is associated with severe toxicity at any higher weight-based doses

- High-level alerts, where the user is not permitted to continue with the prescription ('hard stop') and prescription is thus disallowed. We used these as a surrogate for major prescription errors with the potential to cause serious harm.

The system also prompts doctors to respond to selected *abnormal laboratory test results*, as follows:

- Warnings: Most abnormal laboratory results generate no warning, but more seriously abnormal values, for example, a potassium concentration between 5.5 and 6.5 mmol/L, produces a *warning* when the doctor logs into a patient's record;
- Alarms: The most serious abnormal values, such as a potassium concentration above 6.5 mmol/L, produce an interruptive *alarm* whenever doctors enter the electronic system (regardless of which patient they are viewing) to prompt them to affirm that they have noted and reacted to the specific abnormal result.

The system contains 415 distinct laboratory result warnings and 77 distinct alarms. When responding to them, doctors can either click a button on the electronic system to 'accept' the message (and thus show explicitly that they have acknowledged the clinical implications of the decision to proceed) or click a button to 'ignore' the message. PICS has a comprehensive audit database of actions, including all prescriptions, messages seen by doctors and responses to warnings or alarms.

In UK hospitals, most prescribing is done by junior doctors below the grade of registrar, and they were therefore chosen as the population of interest for this study. We used the database to evaluate associations, for each doctor, between numbers of low-level and high-level prescribing alerts, and between the numbers of intermediate and high-level alerts. Response to prescribing alerts was classified as either 'heeding', where a prescription was altered or abandoned so that the alert was no longer generated, or 'overriding', where the prescription proceeded to completion despite the alert. Similarly, response to laboratory messages was classified as either 'accepting' or 'ignoring'.

Permission to perform this evaluation was obtained from the Clinical Governance Support Unit of the University Hospitals Birmingham NHS Foundation Trust, which deemed this study to be service evaluation not requiring research ethics committee approval.

Data capture

To ensure the study focused on a relatively stable group of doctors over the annual rotation, which begins each year in August, we analysed data from PICS between 8 August 2007 and 31 July 2008. Data were extracted for each junior doctor on all prescriptions generated by PICS, together with information on date and directorate (clinical service division), type and level of any alert, and whether the prescription was completed with or without modification. Only those doctors making more than 20 completed prescriptions during the study period were included (to exclude very short term attachments such as locum doctors prescribing during a few shifts only). All data were fully anonymized.

Information on completed prescriptions was matched to corresponding data on the alerts generated. So too was information on total numbers of completed prescriptions made by each doctor within each directorate. Where matching was not possible, data on such alerts or prescriptions were excluded. The number of laboratory warnings and alarms posted to the doctor and associated responses (accepting/ignoring) were also extracted from the database. All inpatient hospital directorates were included, with the exception of the oncology directorate, where PICS was not fully operational at the time.

Analysis

Rates of prescribing alerts for every 1000 completed prescriptions were calculated for each doctor, by grade of alert, type of warning (e.g. drug–drug interaction), and directorate. The slopes (with standard error and P value) and correlation coefficients for associations between rates of high-level alerts and rates of intermediate and of low-level alerts were evaluated using generalized linear models, weighted by the number of completed prescriptions made. Similar methods were used to measure associations between rates of high-level alerts and propensity to override

intermediate alerts and low-level alerts. These methods were also used to measure association between prescribing alerts and ignoring a warning or alarm triggered by an abnormal laboratory result. To reduce errors due to multiple comparisons, *P* values of less than 0.01 were taken to indicate a statistically significant association.

Funnel plots of rates of high-level alerts against number of completed prescriptions were evaluated with 95% and 99% confidence bands in order to explore heterogeneity between doctors. Prescribers outside these confidence bands may be thought of as outliers, in the sense that they are at the extreme of what would be considered a normal high-level alert rate. To further explore associations between high-level alerts and intermediate or low-level alerts, indicator variables for doctors falling outside the 99% confidence bands for intermediate and low-level alerting rates were superimposed on the funnel plots. Results are presented for all directorates combined and for the three directorates where most prescriptions were completed.

Results

Altogether 849,153 completed prescriptions were issued and 1,094,693 prescribing alerts were generated by PICS during the one year study period. In total, 432 junior doctors used the system (and made 70.8% of all prescriptions in the Trust over this period), but data relating to 50 doctors who completed fewer than 20 prescriptions each over the study period were excluded, along with their 369 prescriptions and 419 alerts. Matching of one further doctor, 93 prescriptions, and 14,856 alerts proved technically impossible and related data were excluded. In addition, 13 prescriptions (associated with 15 alerts) that erroneously entered the database from the oncology directorate were excluded. Available for analysis therefore were 1,079,403 alerts relating to 381 doctors, who completed 848,678 prescriptions over the study period.

Prescribing alerts generated

Most (895,029; 83%) of the 1,079,403 prescribing alerts generated by PICS were low-level alerts.

Table 2
Alerts by type of anomaly

	Alerts (n)			*Heeded alerts* (n)			*Heeded alerts* (%)		
	High-level	*Intermediate*	*Low-level*	*High-level*	*Intermediate*	*Low-level*	*High-level*	*Intermediate*	*Low-level*
Contraindications	908	4535	112,532	908	1443	23,291	100	32	21
Dose	10,529	57,807	552,806	10,529	26,965	249,503	100	47	45
Interaction	503	110,092	229,691	503	9599	29,284	100	9	13
All	11,940	172,434	895,029	11,940	38,007	302,078	100	22	34

Table 3
Alerts by directorate: number, rate and percentage heeded (ranked by high-level alerting rate)

	Alerts (n)				Alert rate per 1000 scripts			Alerts heeded (%)		
Directorate	*High-level*	*Intermediate*	*Low-level*	*Scripts* (n)	*High-level*	*Intermediate*	*Low-level*	*High-level*	*Intermediate*	*Low-level*
Critical Care	1571	16,604	64,467	45,934	34	361	1403	100	28	37
Burns Surgery	155	3328	7277	6791	23	490	1072	100	16	34
Trauma / Orthopaedics	1012	20,252	49,238	58,247	17	348	845	100	11	31
Maxillofacial surgery	242	2810	13,503	14,359	17	196	940	100	20	38
Plastics	315	4707	15,261	20,115	16	234	759	100	15	26
Liver	798	10,794	50,866	49,165	16	220	1035	100	15	31
Neurosciences	688	9352	40,634	46,367	15	202	876	100	20	34
Urology	437	5516	26,079	29,889	15	185	873	100	17	31
Surgery	1476	25,176	98,790	105,045	14	240	940	100	16	34
Medicine	3393	40,575	285,252	274,774	12	148	1038	100	27	35
Ear Nose Throat	236	2815	16,084	19,226	12	146	837	100	20	32
Haematology	268	3365	36,225	24,962	11	135	1451	100	18	34
Cardiothoracic	253	4993	34,288	24,197	10	206	1417	100	23	24
Renal	493	6587	71,858	57,443	9	115	1251	100	25	32
Vascular surgery	175	5602	24,471	20,770	8	270	1178	100	14	32
Cardiology	428	9958	60,736	51,394	8	194	1182	100	49	38
All directorates	11,940	172,434	895,029	848,678						

Fewer (172,424; 16%) were intermediate alerts. A very few (11,940; 1%) were high-level alerts indicative of a serious prescribing error.

Rates of heeding prescribing alerts

Of the 1,079,403 prescribing alerts, 352,025 (33%) overall were heeded (i.e. the prescription was abandoned or changed so that an alert was no longer generated). The remaining alerts were overridden. Of the 172,434 intermediate alerts, 38,007 (22%) were heeded, while of the 895,029 low-level alerts, 302,078 (34%) were heeded. All 11,940 high-level alerts had to be heeded as the alert could not be overridden ('hard stop' warnings).

Issues generating prescribing alerts

Of all prescribing alerts, 621,142 (58%) related to dose-range anomalies, 340,286 (31%) to drug–drug interaction messages, and 117,975 (11%) to contraindications (Table 2). For dose-range anomalies, 47% (99%CI 46.1–47.2%) of intermediate alerts and 45% (99%CI 45.0–45.3%) of low-level alerts were heeded. For contraindications, 32% (99%CI 30.0–33.6%) of intermediate alerts and 21% (99%CI 20.4–21.0%) of low-level alerts were heeded. For drug interactions, doctors were less likely to heed intermediate alerts (8.7% heeded; 99%CI 8.5–8.9%) than low-level alerts (12.8%, 99%CI 12.6–12.9%).

High-level alerts accounted for only a very small proportion of each category: 908/117,975 (0.76%) of all alerts for contraindications; 503/340,286 (0.15%) for drug interactions; and 10,529/621,142 (1.7%) for dose-range anomalies.

Laboratory test result warnings and alarms

Doctors failed to acknowledge 206,580 of the 342,929 warnings (median percentage ignored per doctor 57%; IQR 28–88%) and 49,746 of the 60,062 alarms (median 90%; IQR 70–100%) relating to abnormal laboratory results.

Variations by directorate

The number of completed prescriptions varied considerably by directorate. For example, 274,774

completed prescriptions were issued over the period in the general medical directorate, compared with just 6,791 in the burns surgical directorate (Table 3). The rate of prescribing alerts per 1000 completed prescriptions also varied between directorates. For example, there were eight high-level alerts for every 1000 completed prescriptions in the vascular surgery directorate, compared with 34 per 1000 in the critical care directorate. Rates of low-level alerts varied from 837 per 1000 completed prescriptions in the ear, nose and throat directorate to 1451 in haematology. Rates of heeding of prescription warnings also varied by directorate, for example, 24% of 34,288 low-level warnings generated in the cardiothoracic surgical directorate were accepted, compared with 38% of 60,736 warnings generated in the cardiology medical directorate.

Variation among doctors

The median number of low-level alerts per 1000 completed prescriptions per junior doctor was 1033 (IQR 903–1205), of which 34% (IQR 31–39) were heeded. The median number of intermediate alerts per junior doctor was lower, at 196 (IQR 159–266), of which 23% (IQR 16–30) were heeded. The median number of high-level alerts

Figure 1

Funnel plots of rates of high-level (hard stop) alerts against number of prescriptions. Rates of high-level (hard stop) alerts per 1000 completed prescriptions with mean (solid line) and 95 (dashed lines) and 99 (dotted lines) percent confidence bands. Doctors whose rate of intermediate (password) alerts exceeds the corresponding 95% confidence band are marked with 'x' and doctors whose rate of low-level (tickbox) alerts exceeds the 95% confidence band are marked with a solid dot. a. All Directorates. b. General Surgery Directorate. c. Trauma/Orthopaedic Directorate. d. General Medicine Directorate

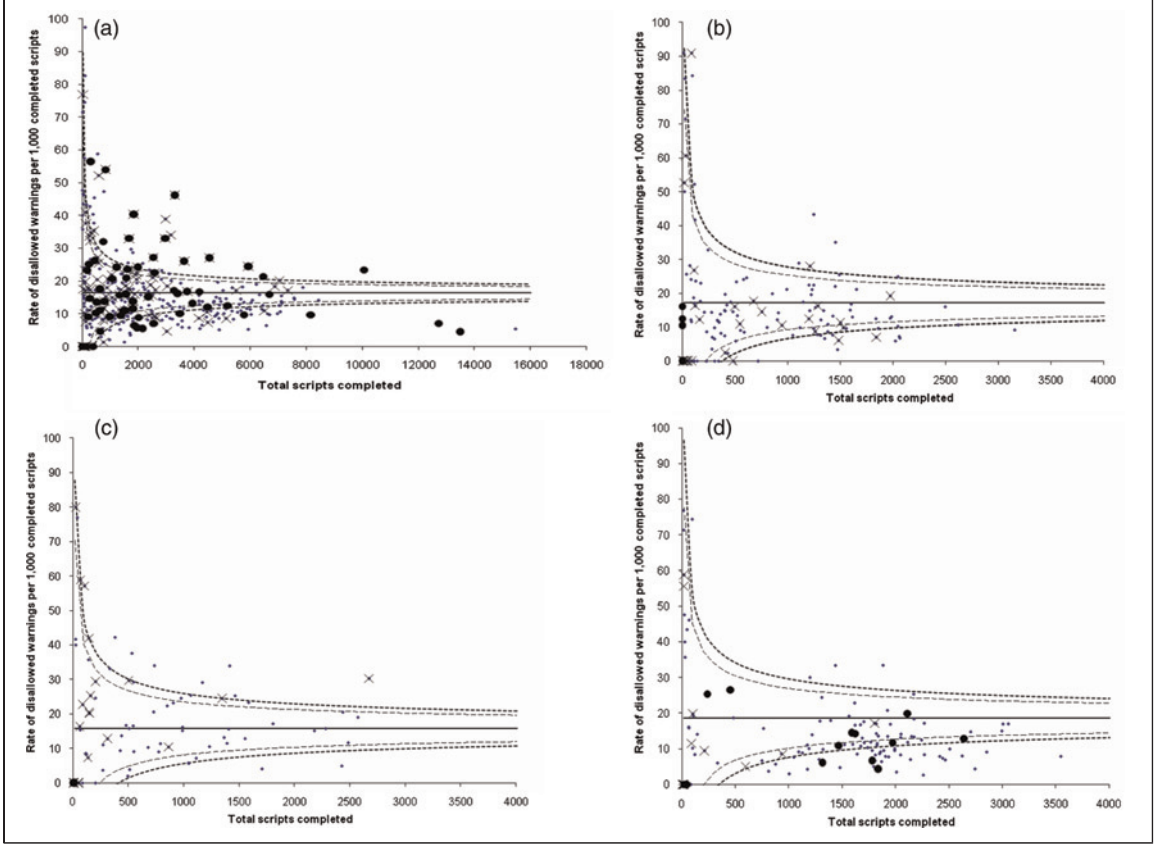

Table 4
Association between rates of hard stop alerts and (a) intermediate and low-level alerts, (b) intermediate and low-level alerts that were not heeded and (c) laboratory warnings and alarms not heeded

	All directorates			*Surgery Directorate*			*Trauma and Orthopaedics Directorate*			*Medical Directorate*		
	Slope (SE)	*P value*	*Correlation*	*Slope (SE)*	*P value*	*Correlation*	*Slope (SE)*	*P value*	*Correlation*	*Slope (SE)*	*P value*	*Correlation*
Rate intermediate alerts	0.045 (0.005)	***0.000***	0.40	0.016 (0.007)	**0.030**	0.15	0.015 (0.006)	***0.001***	0.19	0.022 (0.011)	**0.046**	0.12
Rate low-level alerts	0.005 (0.002)	**0.019**	0.12	0.006 (0.004)	0.138	0.10	0.006 (0.004)	0.220	0.09	0.002 (0.003)	0.494	0.05
Intermediate alerts not heeded (%)	0.031 (0.043)	0.469	0.04	−0.543 (0.188)	***0.004***	−0.22	−0.104 (0.150)	0.518	−0.06	0.008 (0.063)	0.904	0.01
Low-level alerts not heeded (%)	−0.318 (0.087)	***0.000***	−0.19	−0.566 (0.210)	***0.008***	−0.19	−0.223 (0.179)	0.214	−0.10	−0.194 (0.105)	0.066	−0.11
Laboratory warnings not heeded (%)	−0.002 (0.001)	0.226	−0.06	−0.009 (0.010)	0.377	−0.06	−0.000 (0.003)	0.882	−0.01	−0.002 (0.001)	0.324	−0.06
Laboratory alarms not heeded (%)	−0.004 (0.002)	**0.044**	−0.11	−0.016 (0.012)	0.193	−0.09	−0.007 (0.004)	0.091	−0.13	−0.005 (0.002)	***0.005***	−0.18

P values significant at the 99% level are in bold italics
P values significant at the 95% level are in bold

(resulting in the prescription being disallowed) was 13 per 1000 prescriptions (IQR 9–20).

A funnel plot of rate of high-level alerts against the number of completed prescriptions reveals considerable heterogeneity in rates of prescribing alerts between doctors (Figure 1). Outliers on the high-level alerts funnel plot were not consistently identified as outliers for intermediate or low-level alerts (Figure 1a). When stratified by directorate, funnel plots showed less heterogeneity and fewer outliers (Figure 1b–d), but again outliers for high-level alerts were not consistently identified as outliers for rates of low-level or intermediate alert.

Rates of intermediate alerts were statistically associated with the rate of high-level alerts (Table 4), but the correlation was weak ($r = 0.40$, $P = 0.000$), and the correlation between rates of low-level and high-level alerts was even weaker ($r = 0.12$, $P = 0.019$). Correlations between rates of high-level alerts and rates of not heeding low-level or intermediate alerts were also low, (or even negative), ranging from 0.04 to −0.22 across directorates.

Correlations between rates of failure to heed intermediate alerts and failure to heed low-level alerts were weak (ranging from 0.21 to 0.50 between directorates) (Table 5). Not heeding prescription-associated alerts correlated poorly, and sometimes negatively, with ignoring laboratory warnings or alarms (Table 5). Analysis of prescribing alerts by category – dose-range anomaly, interaction and contraindication – gave broadly similar findings (Table 6).

Discussion

This single NHS foundation trust generated over a million prescribing alerts in the one-year period. Only 1% of prescriptions produced high-level alerts, but this amounts to over 10,000 such instances per year. If high-level alerts are reasonable surrogates for serious prescribing errors, then error in hospitals is frequent, as others have also found.[12,13] In targeting serious errors, it would of course be useful if it were possible to screen or monitor individuals most at risk of making a catastrophic prescribing error. Our data suggest that some doctors trigger intermediate alerts more frequently than others, and some

Table 5
Association between not heeding prescription alerts and not heeding laboratory warnings and alarms

	All directorates			*Surgery Directorate*			*Trauma and Orthopaedics Directorate*			*Medical Directorate*		
	Slope (SE)	*P value*	*Correlation*	*Slope (SE)*	*P value*	*Correlation*	*Slope (SE)*	*P value*	*Correlation*	*Slope (SE)*	*P value*	*Correlation*
Associations with ignoring intermediate alerts												
Low-level prescribing alerts not heeded (%)	1.041 (0.094)	***0.000***	0.50	0.313 (0.111)	***0.005***	0.21	0.429 (0.104)	***0.000***	0.37	0.682 (0.106)	***0.000***	0.40
Laboratory warnings not heeded (%)	0.003 (0.002)	0.080	0.09	0.002 (0.002)	0.239	0.09	−0.007 (0.002)	***0.000***	−0.34	−0.001 (0.002)	0.712	−0.02
Laboratory alarms not heeded (%)	−0.001 (0.002)	0.816	−0.01	−0.000 (0.002)	0.936	−0.00	−0.007 (0.003)	**0.019**	−0.23	0.005 (0.002)	**0.012**	0.18
Associations with ignoring low-level alerts												
Laboratory warnings not heeded (%)	0.002 (0.001)	***0.003***	0.15	0.000 (0.001)	0.497	0.05	−0.000 (0.001)	0.954	−0.03	0.002 (0.000)	**0.021**	0.14
Laboratory alarms not heeded (%)	0.001 (0.002)	0.675	0.02	0.001 (0.001)	0.282	0.08	−0.000 (0.002)	0.790	−0.02	0.002 (0.001)	0.090	0.11

P values significant at the 99% level are in bold italics
P values significant at the 95% level are in bold
Slope is the coefficient from a fitted generalized linear model, modelling the linear association between rate of hard stop alerts and each of the listed independent variables within the table

doctors trigger high-level alerts more frequently than others, but these 'outliers' tend not to be the same doctors, and nor is there a relationship between tendency to provoke alerts or warnings/alarms and 'heeding' behaviour in response. Our analysis suggests that it is not possible to use routine prescribing data recording behaviour in relation to alerts, warnings and alarms to identify doctors who are more likely to generate an alert indicative of a serious prescribing error.

Our study does have a number of limitations. It was conducted in a single NHS Teaching Hospital Trust, using a specific computer system. Other IT systems might produce different findings, and it may be useful to replicate the same study in other systems and to use methods of triangulation to assess more holistically issues of individual variation in prescribing behaviour.

Our data reveal considerable variation between doctors in rates of high-level alerts: some doctors generate such alerts frequently, while others do so infrequently. Similarly, doctors vary widely in rates of low and intermediate level alerts they generate, and in rates of generated alerts that they heed. The number of high-level alerts a doctor generated correlates only weakly with the number of intermediate alerts, or the number of low-level alerts the same doctor generates. Furthermore, there is little or no correlation between the number of prescribing alerts of any grade a doctor generates, and the doctor's propensity to *heed* intermediate or low-level alerts. This means that doctors who are at highest risk of making serious prescribing errors, as reflected by triggering high-level alerts, cannot be identified from the rate of low-level or intermediate alerts they generate, or from whether they heed them. One interpretation is that the search for the phenotype of a generally error-prone person may prove as elusive in medicine as it has been elsewhere,[14–19] but the available data do not allow more certain conclusions to be drawn.

More generally our analysis demonstrates some of the limitations of using routine data from a computerized prescribing system as a way of detecting 'error' or aspects of individual performance. The extent to which different levels of alert are reasonable surrogates for varying gravity of error is, for example, poorly understood. Doctors routinely override prescribing alerts[20] as most can be safely ignored. Our analysis suggests that there are clear risks of using electronic traces of

Table 6
Association between rates of high-level (hard stop) alerts and intermediate and low-level alerts by category of warning

	Interactions			*Dose*			*Contraindications*		
	Slope (SE)	*P value*	*Correlation*	*Slope (SE)*	*P value*	*Correlation*	*Slope (SE)*	*P value*	*Correlation*
Rate intermediate alerts	0.000 (0.001)	***0.000***	0.23	0.038 (0.010)	***0.000***	0.19	−0.027 (0.015)	0.080	−0.09
Rate low-level alerts	−0.001 (0.000)	0.079	−0.09	0.015 (0.003)	***0.000***	0.29	−0.000 (0.001)	0.806	−0.01
Intermediate alerts not heeded (%)	0.008 (0.008)	**0.021**	0.12	−0.010 (0.038)	0.802	−0.01	−0.000 (0.003)	0.748	−0.02
Low-level alerts not heeded (%)	0.013 (0.008)	**0.021**	0.10	−0.108 (0.081)	0.182	−0.07	−0.012 (0.008)	0.105	−0.08
Laboratory warnings not heeded (%)	−0.000 (0.000)	**0.028**	−0.11	−0.001 (0.001)	0.359	−0.05	−0.000 (0.000)	0.342	−0.05
Laboratory alarms not heeded (%)	−0.000 (0.000)	0.135	−0.08	−0.003 (0.002)	0.053	−0.11	−0.000 (0.000)	0.654	−0.02

P values significant at the 99% level are in bold italics
P values significant at the 95% level are in bold
Slope is the coefficient from a fitted generalized linear model, modelling the linear association between rate of hard stop alerts and each of the listed independent variables within the table

rule-breaking (such as apparent breaches of protocols or overrides of warning recorded on IT systems) as an indication that something that might injure a patient has occurred. For example, we found no association between doctors breaking the rule that a box should be ticked to confirm that an abnormal laboratory result has been noted and the chance that they will make a serious prescribing error indicated by a 'hard stop' on the computer system. Here not only is there no positive correlation, but the correlations turn negative. Such an observation supports the possibility that some rule-following may be little more than a ritualised display of compliance rather more than a signal of safe practice. This finding points to the possible risks of using data from hospital IT systems to try to alter behaviour. For example, our data provide some indications that encouraging doctors to respond to low-level computer-generated warnings will have little impact on major errors, and risks diverting efforts from preventing low-frequency high-harm events to events that are high-frequency but low-harm or no-harm. Further, encouraging staff to respond to such prompts, which they may see as purely bureaucratic, could damage the overall legitimacy of the patient safety enterprise.

IT systems are widely and increasingly used to collect routine data in healthcare organizations. Organizations may see these data as a potential source of quality assurance information, but need to consider what criteria need to be met before acting on it. Evaluations of what kinds of questions such systems should be used for will become increasingly valuable. There is little evidence, on the basis of our analysis, that routine prescribing systems can or should be used to detect doctors who are more likely to make a serious prescribing error. Though caution is required in this particular application, use of routine data from such systems may, nonetheless, have an important role in monitoring quality in healthcare organizations, and future studies should investigate their potential in this area.

References

1 Powell AE, Davies HTO, Thomson RG. Using routine comparative data to assess the quality of healthcare: understanding and avoiding common pitfalls. *Qual Safety Health Care* 2003;**12**:122–8
2 Marshall MN, Shekelle P, Davies HTO, Smith PC. Public reporting on quality in the United States and the United Kingdom. *Health Affairs* 2003;**22**:134–48

3 Davies H. *Measuring and reporting the quality of care: issues and evidence from the international research literature*. Edinburgh: NHS Quality Improvement Scotland, 2006. See http://www.nhsqis.org.uk/nhsqis/files/Davies%20Paper.pdf

4 McManus C, Vincent C. Selecting and educating safer doctors. In: Vincent C, Ennis M, Audley R, eds. *Medical Accidents.* Oxford: Oxford University Press, 2003

5 Dornan T, Ashcroft D, Heathfield H, *et al.* An in depth investigation into causes of prescribing errors by foundation trainees in relation to their medical education – EQUIP Study. London: General Medical Council, 2009

6 Leape LL, Brennan TA, Laird N, *et al.* The nature of adverse events in hospitalized patients. Results of the Harvard medical practice study II. *N Engl J Med* 1991;**324**:377–84

7 Pirmohamed M, James S, Meakin S, *et al.* Adverse drug reactions as cause of admission to hospital: prospective analysis of 18 820 patients. *BMJ* 2004;**329**:15–19

8 Dobrzanski S, Hammond I, Khan G, Holdsworth H. The nature of hospital prescribing errors. *Brit J Clin Govern* 2002;**7**:187–93

9 Ferner RE. The epidemiology of medication errors: the methodological difficulties. *Brit J Clin Pharmacol* 2009;**67**:614–20

10 Nightingale PG, Adu D, Richards NT, Peters M. Implementation of rules based computerised bedside prescribing and administration: intervention study. *BMJ* 2000;**320**:750–3

11 British Medical Association/Royal Pharmaceutical Society of Great Britain. *British National Formulary No. 60*. London: BMA, 2010

12 Lin CP, Payne TH, Nichol WP, Hoey PJ, Anderson CL, Gennari JH. Evaluating clinical decision support systems: monitoring CPOE order check override rates in the Department of Veterans Affairs' Computerized Patient Record System. *J Am Med Inform Assoc* 2008;**15**:620–6

13 Van Der Sijs H, Aarts J, Vulto A, Berg M. Overriding of drug safety alerts in computerized physician order entry. *J Am Med Inform Assoc* 2006;**13**:138–47

14 Lawton R, Parker D. Individual differences in accident liability: A review and integrative approach. *Hum Factors* 1998;**40**:655–71

15 Lawton R, Parker D, Manstead ASR, Stradling SG. The role of affect in predicting social behaviours: the case of road traffic violations. *J Appl Soc Psychol* 1997;**27**:1258–76

16 Booysen AE, Erasmus JAK. The relationship between some personality factors and accident risk. *S Afr J Psychol* 1989;**19**:144–52

17 Hilakivi I, Veilahti J, Asplund P, Sinivuo J, Laitinen L, Koskenvuo K. A 16-factor personality test for predicting automobile driving accidents of young drivers. *Accident Anal Prev* 1989;**21**:413–18

18 Elander J, West R, French D. Behavioural-correlates of individual-differences in road-traffic crash risk – an examination of methods and findings. *Psychol Bull* 1993;**113**:279–94

19 Lardent CL. Pilots who crash – personality constructs underlying accident prone behaviour of fighter pilots. *Multivar Exp Clin Res* 1991;**10**:1–25

20 Isaac T, Weissman JS, Davis RB, *et al.* Overrides of medication alerts in ambulatory care. *Arch Intern Med* 2009;**169**:305–11