Costs and Cost-Effectiveness of User-Testing of Health Professionals’ Guidelines to Reduce the Frequency of Intravenous Medicines Administration Errors by Nurses in the United Kingdom: A Probabilistic Model Based on Voriconazole Administration

Matthew D. Jones1 · Bryony Dean Franklin2,3 · D. K. Raynor4,5 · Howard Thom6 · Margaret C. Watson7 · Rebecca Kandiyali6

Accepted: 12 July 2021 © The Author(s) 2021

Abstract

Aim In the UK, injectable medicines are often prepared and administered by nurses following the Injectable Medicines Guide (IMG). Our earlier study confirmed a higher frequency of correct administration with user-tested versus standard IMG guidelines. This current study aimed to model the cost-effectiveness of user-testing.

Methods The costs and cost-effectiveness of user-testing were explored by modifying an existing probabilistic decision-analytic model. The adapted model considered administration of intravenous voriconazole to hospital inpatients by nurses. It included 11 error types, their probability of detection and level of harm. Model inputs (including costs) were derived from our previous study and other published data. Monte Carlo simulation using 20,000 samples (sufficient for convergence) was performed with a 5-year time horizon from the perspective of the 121 NHS trusts and health boards that use the IMG. Sensitivity analyses were undertaken for the risk of a medication error and other sources of uncertainty.

Results The net monetary benefit at £20,000/quality-adjusted life year was £3,190,064 (95% credible interval (CrI): −346,709 to 8,480,665), favouring user-testing with a 96% chance of cost-effectiveness. Incremental cost-savings were £240,943 (95% CrI 43,527–491,576), also favouring user-tested guidelines with a 99% chance of cost-saving. The total user testing cost was £6317 (95% CrI 6012–6627). These findings were robust to assumptions about a range of input parameters, but greater uncertainty was seen with a lower medication error risk.

Conclusions User-testing of injectable medicines guidelines is a low-cost intervention that is highly likely to be cost-effective, especially for high-risk medicines.

Key Points for Decision Makers

User-testing injectable medicines guidelines for nurses reduces the number of medication errors made when giving intravenous medicines to hospital inpatients.

As a result, the cost of the user-testing process is likely to be outweighed by the increased health of patients and reduced health service costs for treating the harm caused by medication errors.

Therefore, user-testing of guidelines for injectable medicines is a relatively low-cost and cost-effective intervention.
1 Introduction

Medication errors are a leading cause of avoidable patient harm, costing an estimated $42 billion per annum worldwide [1]. Each year in the USA, there are an estimated 1.2 million hospitalisations affected by an injectable medicine error, increasing costs by $2.7–5.1 billion [2]. Errors are particularly common with intravenous medicines, with 35–48% of doses containing at least one error [3, 4]. In UK observational studies, 22–74% of adult intravenous medication errors have been classified as having moderate to severe potential consequences [5–7]. In a point prevalence study, 90% of intravenous medication errors were classified as unlikely to cause harm, but 9% were likely to require increased monitoring or intervention and 0.4% to cause temporary harm [8]. Of the many potential causes of patient safety incidents, poor quality written guidance for health professionals has received little attention from researchers [9, 10].

In the UK, nurses prepare and administer most intravenous doses on hospital wards. In over 120 hospital organisations, they use the NHS (National Health Service) Injectable Medicines Guide (IMG) website as a source of written information on how to do this for each dose. Two linked studies have shown that nurses using a user-tested version of the IMG were more likely to correctly prepare and administer an intravenous medicine, and did this more quickly, than nurses using the current version [11, 12]. User-testing is a process based on iterative rounds of interviews with users to discover where they misunderstand a written document and to test potential improvements [11]. This process was applied to the current IMG guideline for voriconazole, which was chosen as it is a high-risk medicine requiring multiple stages of preparation [11]. This resulted in a user-tested voriconazole guideline. In our in situ simulation experiment (hereafter described as the “clinical study”), 273 hospital nurses were randomised to use existing or user-tested versions of the guidelines to administer an infusion of voriconazole to a manikin arm while working on their usual ward [12]. Direct observation was used to identify 11 different types of medication error (Table A of the Online Supplemental Material), which were defined as any deviation from the simulated medication order, the hospital’s policies or the IMG. The potential level of harm for each observed error was assigned using a validated method by seven physicians, pharmacists and nurses who independently scored each error on a scale from 0 (no harm) to 10 (death). A mean score of < 3 was classified as minor harm, a mean score of 3–7 was moderate harm and a mean score > 7 was severe harm. More infusions were completed without any IMG-related errors with the user-tested guidelines (n = 67, 48%) than with current guidelines (n = 26, 20%) (risk ratio 2.46; 95% confidence interval (CI) 1.68–3.60). As medication errors can cause harm and thus increase health service costs [13], there is potential for the costs of user-testing to be offset by cost-savings resulting from fewer errors. However, the clinical study did not consider cost-effectiveness, so it is unclear whether user-testing of IMG guidelines should be introduced.

Economic decision analytic models can synthesise and extrapolate evidence from randomised trials and other sources to compare the expected costs and consequences of different decision options [14]. The aim of this study, therefore, was to build an economic decision model to assess the cost-effectiveness of user-testing injectable medicines guidelines. To date, studies of the economic impact of medication errors have largely considered their effects on health service costs [15, 16]. Similarly, many economic evaluations of interventions to prevent medication errors have considered only the cost of the intervention and the health services costs avoided by reducing errors [17, 18]. Few of these studies have also considered patients’ quality of life, although this has been recently recommended [15]. Such an approach will enable a more complete assessment of the effects of error-prevention interventions and enable direct comparison with the cost-effectiveness of other interventions. It may also increase uncertainty due to the limitations of the available data on the effects of errors on quality of life. This study therefore had two specific objectives: (i) to investigate the cost impact of user-testing, by comparing the costs of the user-testing process with the reduced costs resulting from fewer medication errors and faster drug preparation, and (ii) to investigate the cost-effectiveness of user-testing in terms of net monetary benefits (derived from quality-adjusted life years; QALYs).

2 Methods

There are numerous types of medication error and therefore many potential pathways for the development of harmful effects. Economic models of interventions to reduce medication errors therefore usually consider general categories of harm severity, rather than specific pathways. To do this, several parameters must be determined, including the number of errors with and without the intervention, the likelihood of detection of errors before administration, the level of harm produced, the cost of treating such adverse effects, and the intervention costs. In addition, to achieve our second objective, we also needed to consider the effects of harmful errors of different levels of severity on patients’ quality of life.

Three previous economic models of interventions to reduce medication errors were identified (Table 1). The probabilistic decision tree model developed by Karnon et al.
Cost-Effectiveness of User-Testing Guidelines

Table 1  Summary of previous economic models of interventions to reduce medication errors

| First author and year | Study population and setting | Modelling method | Categorisation of harm from errors | Outcome measures |
|-----------------------|-----------------------------|-----------------|-----------------------------------|-----------------|
| Karnon, (2008) [21]   | Inpatients of a 400-bed acute hospital in the UK | Decision tree | No harm, minor, moderate and severe/life-threatening harm | Incremental costs of correcting an error detected before administration, and treating the adverse effects of an error including extra tests, treatments and hospital length of stay; Monetary valuation of QoL reduction following a harmful error |
| Lahue, (2012) [2]     | Hospital inpatients receiving injectable medicines in the USA | Not specifically named, but effectively a decision tree | Harm and no harm | Incremental cost of a harmful medication error |
| Samp, (2014) [22]     | Primary, secondary and tertiary care patients in the USA | Decision tree | 8 categories from combining: temporary, permanent or no harm; hospitalisation; treatment required; death | Incremental costs of the error, including monitoring and those arising from permanent harm |

QoL, quality of life

[19] most closely meets the requirements of our objectives, as it relates to errors involving medication use in a UK hospital, considers quality of life, and classifies the harm that may follow a medication error into similar categories to those used in our previous clinical study. We therefore chose to adapt this model to compare the costs and cost-effectiveness of user-tested injectable medicines guidelines compared with current guidelines.

We followed the Consolidated Health Economic Evaluation Reports Standards [20].

2.1 Perspective, Population and Setting

The IMG is currently used by nurses preparing medicines in the hospitals of 121 NHS trusts and health boards throughout the UK (Robin Burfield, NHS Wales Informatics Service, personal communication, 30 July 2019). These organisations also share resources to maintain and develop the guide. We therefore performed analysis from the perspective of the NHS payer. The population was hospital inpatients receiving intravenous voriconazole in all 121 hospital trusts and health boards within our perspective (see Sect. 2.6.4).

2.2 Intervention and Comparator

The intervention was the user-testing process to revise the design of the IMG [11]. This involved four pilot interviews followed by three iterative rounds of interviews each with ten nurses, which identified problems finding and understanding information. Each round was followed by revisions to the IMG to resolve problems. We assumed that user-tested guidelines would be introduced by updating the IMG website, with the same cost as provision of the current version of the IMG (Sect. 2.7.3). We compared use of the user-tested IMG with the current version of the IMG.

2.3 Model Outcomes

All outcomes were calculated across the time horizon (Sect. 2.4) for the whole cohort of patients from 121 NHS trusts and health boards within our perspective (Sect. 2.1). We calculated the mean decrease in the total number of medication errors that cause moderate-severe patient harm (i.e. preventable adverse drug events; pADEs) with the user-tested IMG (the primary outcome of our previous clinical study [12]). We also calculated the mean decrease in the total number of pADEs with the user-tested IMG. These were calculated by extrapolating the frequency of medication errors observed with the current and user-tested IMG in our previous clinical study through our model to the total number of doses of voriconazole given in all 121 NHS organisations within our perspective (Sects 2.5 and 2.6).

We calculated the incremental cost-saving (ICS) as the difference between health service costs with the current IMG and health service costs with the user-tested IMG. This included the costs associated with medicines administration (nursing time), medication errors (correcting an error detected before administration, and treating the adverse effects of an error including extra tests, treatments and hospital length of stay), and the user-testing intervention. Due to the complexity of including insured costs, we did not consider litigation costs as they represent only 0.5% of the total cost of NHS medication errors [16].

△ Adis
M. D. Jones et al.

We assumed that quality of life would be decreased by a medication error, so we assessed the benefits of user-testing using QALY decrements following a medication error. User-tested guidelines would therefore be more beneficial than current guidelines if they result in a smaller QALY decrement. We evaluated cost-effectiveness by calculating net monetary benefit (NMB). NMB represents the additional value in monetary terms generated by user-testing, so a positive NMB would indicate that user-testing is optimal. The NMB was calculated as the difference between expected incremental benefits (QALY decrements valued at a willingness-to-pay threshold of £20,000 per QALY) and expected incremental costs. We assessed uncertainty in the user-testing decision by reporting 95% credible intervals and the probability of a positive ICS or NMB, as well as cost-effectiveness acceptability curves.

2.4 Time Horizon and Discounting

We used a 5-year time horizon. As user-testing results in changes to the design of the IMG, its effects persist and further user-testing is not required unless there are changes in practice or training. We considered 1- and 10-year time horizons in sensitivity analyses. We assumed that user-testing took place once at the start of the time horizon, so user-testing costs were not discounted. An annual discount rate of 3.5% was applied to other costs and QALYs [23].

2.5 Model Structure

The Karnon et al. decision tree model [21] considers six types of decision node in the following order: presence or absence of a prescribing error; a dispensing error or an administration error; the error type (wrong drug, dose, route or frequency); detection before administration; resultant level of harm. The present study relates only to medication administration, and we assumed that costs and QALYs associated with prescribing and dispensing are the same with user-tested and current guidelines. Therefore, we adapted the Karnon model by removing the prescribing and dispensing error nodes. The four error types in the Karnon model do not describe the range of errors observed in our clinical study, so we replaced them with the 11 types of IMG-related error from our previous study (Table A of the Online Supplemental Material) [12]. The resultant decision model used for the present study is shown in Fig. 1.

2.6 Model Inputs to Estimate the Number of Medication Errors and Preventable Adverse Drug Events (pADEs)

Data sources used to populate the model to estimate the number of medication errors and pADEs are described in the following sections. Input parameter values and distributions for probabilistic analysis are presented in Table 2. We populated two versions of the model with different probabilities.
| Variable                                                                 | Estimates (95% confidence interval)                        | Distribution                                                                 | Source                                                                 |
|-------------------------------------------------------------------------|------------------------------------------------------------|-----------------------------------------------------------------------------|------------------------------------------------------------------------|
| **Probabilities of error**                                              |                                                            |                                                                             |                                                                        |
| Error type probabilities with current guidelines (node 1, Fig. 1)       | See Table C of Online Supplemental Material                | Dirichlet (no error then error types I to II):                              | Previous clinical study [12], n = 133                                  |
|                                                                         |                                                            | Dirichlet (26.3, 0.3, 0.3, 10.8, 0.3, 6.3, 27.3, 0.3, 3.3, 33.3, 0.3, 27.8)     |                                                                        |
| Error type probabilities with user-tested guidelines (node 1, Fig. 1)   | See Table C of Online Supplemental Material                | Dirichlet (no error then error types I to II):                              | Previous clinical study [12], n = 140                                  |
|                                                                         |                                                            | Dirichlet (67.3, 0.3, 0.3, 4.3, 1.3, 0.3, 48.3, 0.3, 6.3, 11.3, 0.3, 3.3)     |                                                                        |
| Probability an error is undetected prior to administration (node 2, Fig. 1) | 0.775 (0.718–0.837)                                         | Log-normal (-0.25, 0.086)                                                  | Random effects meta-analysis of all three randomised controlled trials [24–26] identified by a recent systematic review of the effectiveness of nurse double-checking of medicines administration [27]. The input distribution was based on the 95% prediction interval. See Online Supplemental Material page 13. n = 89,006 |
|                                                                         | 95% prediction interval: 0.655–0.918                       |                                                                             |                                                                        |
| Probability an undetected error causes no harm (node 3, Fig. 1)         | 0.904 (0.867–0.942)                                         | Beta (217, 23)                                                              | Proportion of observed errors from the ECLIPSE study rated as causing no harm (NCCMERP severity rating C) [8], n = 240 |
| Probabilities of minor, moderate or severe harm for a harmful, undetected error (node 4, Fig. 1) | See Table D of Online Supplemental Material | See Table 3                                                                | Previous clinical study [12]                                           |
| Number of doses of intravenous voriconazole administered using IMG per annum | 4000                                                       | Deterministic                                                              | Voriconazole administration and supply data provided by NHS organisations within our perspective - see Online Supplemental Material page 20 |
| **Medicine administration time**                                        |                                                            |                                                                             |                                                                        |
| Medicine administration time (current guidelines), min                    | 13.3 (12.5–14.2)                                           | Normal (13.3, 0.41)                                                       | Previous clinical study [12], n = 133. Converted to a cost by multiplying by the hourly cost of a band 5 nurse (see below) |
| Medicine administration time (user-tested guidelines), min               | 11.9 (11.3–12.5)                                           | Normal (11.9, 0.31)                                                       | Previous clinical study [12], n = 140. Converted to a cost by multiplying by the hourly cost of a band 5 nurse (see below) |
| **Medication error costs and QALY decrements**                          |                                                            |                                                                             |                                                                        |
| Cost of an error detected before administration                          | £0.27 (0.03–2.39)                                           | Log-normal (−1.29, 1.10)                                                  | Karnon et al. [21] (see Online Supplemental Material pages 21–22)      |
| Treatment cost for a minor pADE                                          | £124 (94–163)                                              | Log-normal (4.82, 0.14)                                                   |                                                                        |
| Treatment cost for a moderate pADE                                        | £1252 (1092–1436)                                          | Log-normal (7.13, 0.07)                                                   |                                                                        |
| Treatment cost for a severe pADE                                         | £1846 (1643–2074)                                          | Log-normal (7.52, 0.06)                                                   |                                                                        |
| QALY decrement following a minor pADE                                     | 0.004 (0.002–0.006)                                        | Normal (0.004, 0.0011)                                                   |                                                                        |
| QALY decrement following a moderate pADE                                  | 0.035 (0.017–0.052)                                        | Normal (0.035, 0.0089)                                                   |                                                                        |
| QALY decrement following a severe pADE                                    | 3.50 (1.87–5.13)                                           | Normal (3.50, 0.83)                                                      |                                                                        |
Table 2 (continued)

| Variable | Estimates (95% confidence interval) | Distribution | Source |
|----------|-------------------------------------|--------------|--------|
| **User-testing resources** | | | |
| Length of pilot user-testing interview, min$^c$ | 24 (19–28) | Normal (24, 2.2)$^a$ | User-testing study [11]. Converted to a cost by multiplying by sum of the hourly costs (see below) of a band 8a pharmacist (user-tester) and a band 5 nurse (participant) |
| Length of round 1 user-testing interview, min$^c$ | 29 (24–33) | Normal (29, 2.5)$^a$ |
| Length of round 2 user-testing interview, min$^c$ | 23 (18–27) | Normal (23, 2.1)$^a$ |
| Length of round 3 user-testing interview, min$^c$ | 20 (17–22) | Normal (20, 1.2)$^a$ |
| Number of pilot interviews | 4 | Deterministic | User-testing study [11] |
| Number of interviews in rounds 1–3 | 10 | Deterministic |
| Interview transcription costs, £/spoken min | 1.75 | Deterministic |
| Length of pilot interview analysis, min$^c$ | 94 (82–106) | Normal (94, 6.1)$^a$ | User-testing study [11]. Converted to a cost by multiplying by hourly cost of a band 8a pharmacist (see below) |
| Length of round 1 interview analysis, min$^c$ | 76 (69–82) | Normal (76, 3.5)$^a$ |
| Length of round 2 interview analysis, min$^c$ | 53 (39–66) | Normal (53, 6.9)$^a$ |
| Length of round 3 interview analysis, min$^c$ | 50 (46–54) | Normal (50, 2.2)$^a$ |
| Time to revise guides after round 1, min$^{cd}$ | 265 (226–304) | Normal (265, 20)$^a$ | User-testing study [11]. Converted to a cost by multiplying by hourly cost of a band 8a pharmacist (see below) |
| Time to revise guides after round 2, min$^{cd}$ | 140 (118–162) | Normal (140, 11)$^a$ |
| Time to revise guides after round 3, min$^{cd}$ | 140 (118–162) | Normal (140, 11)$^a$ |
| User-tester training costs (course cost, trainee time and travel), £$^e$ | 562 | Deterministic | User-testing study [11] |
| Equipment costs, £$^e$ | 4 | Deterministic | User-testing study [11] |

| **Staff costs** | | | |
| Hourly NHS cost of a hospital pharmacist to carry out user-testing, £ | 65 | Deterministic | Personal and Social Services Research Unit data [28] |
| Hourly NHS cost of a hospital nurse as user-testing participant or to administer voriconazole, £ | 37 | Deterministic | Personal and Social Services Research Unit data [28] |

| **Discount rate** | | | |
| Discount rate | – 0.035 | Deterministic | NICE methods of technology appraisal [23]. |

*IMG* Injectable Medicines Guide, *NCCMERP* National Coordinating Centre of Medication Error Reporting and Prevention, *NHS* National Health Service, *NICE* National Institute for Health and Care Excellence, *QALY* quality-adjusted life year, *QoL* quality of life

$^a$Distribution uses standard error

$^b$Distribution uses standard deviation

$^c$Previous user-testing study developed IMG guidelines for two medicines, so times halved

$^d$Assumed relative standard error of 8%, (the mean relative standard error for interview and analysis time)

$^e$Cost assumed to be shared between ten user tests
for each type of error and the level of harm from an undetected error (nodes 1 and 4 in Fig. 1) for current guidelines and user-tested guidelines.

### 2.6.1 Error Type and Level of Harm Caused by a pADE

We used the frequency of errors observed in our previous clinical study (section I.0) [12] to calculate the probabilities of each of the error types or no medication error (node 1, Fig. 1), and for the level of harm caused by a pADE resulting from each error type (node 4, Fig. 1), for both types of guideline. In our previous study, multiple errors were observed in some doses. In line with our previous analysis, in our base case we only considered the error type with the highest potential severity in each observation (Online Supplemental Material, page 3), which we used to define Dirichlet distributions for the error type probabilities (Table 2, applied to node 1 of Fig. 1) and for the probabilities of minor, moderate or severe harm related to each error type (Table 3, applied to node 4 of Fig. 1), for both the current and user-tested guidelines. In the previous clinical study, there were only 273 participants and so rare events (e.g. occurring in fewer than one in 500 doses) were unlikely to have been observed.

Using the data directly would have yielded zero probabilities for such events. We therefore assumed a Dirichlet prior distribution with a count of 0.1 for all error types and severities, and used Bayesian updating with the 273-participant dataset to generate a posterior distribution. This posterior distribution would be guaranteed to have a very low, but non-zero, probability for events that did not take place in the sample.

### Table 3 Base-case distributions for the probabilities of minor, moderate or severe harm for a harmful, undetected error for all error types and with both current and user-tested guidelines

| Error type | Dirichlet distributions (minor harm, moderate harm, severe harm) |
|------------|---------------------------------------------------------------|
|            | Current guidelines                                           | User-tested guidelines |
| I1         | Dirichlet (0.1, 0.1, 0.1)                                    | Dirichlet (0.1, 0.1, 0.1) |
| I2         | Dirichlet (0.1, 0.1, 0.1)                                    | Dirichlet (0.1, 0.1, 0.1) |
| I3         | Dirichlet (5.1, 5.6, 0.1)                                    | Dirichlet (2.1, 2.1, 0.1) |
| I4         | Dirichlet (0.1, 0.1, 0.1)                                    | Dirichlet (0.1, 1.1, 0.1) |
| I5         | Dirichlet (0.1, 6.1, 0.1)                                    | Dirichlet (0.1, 0.1, 0.1) |
| I6         | Dirichlet (0.1, 27.1, 0.1)                                   | Dirichlet (3.1, 45.1, 0.1) |
| I7         | Dirichlet (0.1, 0.1, 0.1)                                    | Dirichlet (0.1, 0.1, 0.1) |
| I8         | Dirichlet (0.1, 3.1, 0.1)                                    | Dirichlet (0.1, 6.1, 0.1) |
| I9         | Dirichlet (4.1, 29.1, 0.1)                                   | Dirichlet (0.1, 11.1, 0.1) |
| I10        | Dirichlet (0.1, 0.1, 0.1)                                    | Dirichlet (0.1, 0.1, 0.1) |
| I11        | Dirichlet (19.1, 3.6, 5.1)                                   | Dirichlet (0.1, 2.1, 1.1) |

These distributions were applied to node 4 of Fig. 1

*a Distribution based entirely on a priori assumption of 0.1 for each error type and severity (see Sect. 2.6.1)

We also conducted a structural sensitivity analysis considering the effects of multiple errors per dose, assuming that their costs and QALY decrements were additive (Online Supplemental Material, pages 6–9).

In the clinical study, 66% of doses contained an error related to the IMG [12], whereas systematic reviews suggest that 35–48% of actual intravenous doses contain an error [3, 4]. Therefore, another sensitivity analysis explored the effects of adjusting the medication error frequency with current guidelines to a rate similar to that reported in practice (32%) while maintaining an unchanged relative reduction in error frequency after user-testing, by adjusting the probabilities associated with nodes 1 and 4 of Fig. 1 (see Tables I–L of the Online Supplemental Material). To explore the consequences of user-testing being less effective when medication errors are less common (i.e. medicines that are simpler to prepare and administer), a further sensitivity analysis considered the relative effects of user-testing being halved when the frequency of medication errors is reduced to 32% of doses (Tables I–L of the Online Supplemental Material).

### 2.6.2 Probability of Error Detection

The Karnon model includes a step accounting for the possible detection of an error before it reaches the patient (node 2, Fig. 1). In this study, the only opportunity to detect an error is if a prepared dose is double-checked by a second nurse, which is recommended practice within the NHS [29]. We therefore used the risk ratio for a medication administration error following a double check by a nurse (compared with no double-check) to populate the model (node 2 of Fig. 1). This combines the probability of a double check being completed and the probability of it detecting an error. A random effects meta-analysis of published data (Online Supplemental Material, page 13) yielded the parameters described in Table 2. The Online Supplemental Material (page 13) describes sensitivity analyses of this risk ratio.

### 2.6.3 Probability that an Undetected Error Causes No Harm

There are few data on the actual consequences of medication errors, as research tends to report the potential consequences had an error not been intercepted. However, not every medication error that reaches a patient produces its potential harm [16, 22]. We specified that for our model this probability should relate to UK intravenous medicines administration, due to the high risk of this route of administration and significant international differences between medication administration systems. A recent systematic review of intravenous medication errors in the UK identified no suitable data [30]. However, a more recent UK study identified intravenous infusion errors using a point prevalence approach, in which errors were identified after administration had commenced.
A total of 240 errors were observed, of which 217 were rated as causing no harm [31]. We therefore populated node 3 of Fig. 1 with the parameters described in Table 2. As these data are derived from only one study, we included both increased (0.99) and decreased (0.75) probabilities in our sensitivity analyses (Table M of the Online Supplemental Material).

### 2.6.4 Number of Doses of Voriconazole Administered Using the Injectable Medicines Guide (IMG) per Annum

We used the number of doses of voriconazole administered using the IMG per annum to convert the probability of each outcome of the decision tree (Fig. 1) into an estimated number of patients. Using voriconazole administration and supply data provided by NHS organisations within our perspective (Online Supplemental Material, page 20), we estimated this value to be between 4130 and 22,980 doses per annum. We used a conservative estimate of 4000 doses in our base-case analysis and carried out a sensitivity analysis with 20,000 doses.

### 2.7 Model Inputs to Estimate Costs and Benefits

We converted the estimated number of medication errors and pADEs into estimated health service costs (see Sect. 2.3) and QALY decrements for a pADE, using the parameters described in Table 2.

#### 2.7.1 Nurse Medicine Administration Time and Costs

Our clinical study measured the time taken to prepare and administer a dose of voriconazole with current and user-tested guidelines [12]. We used these data to define normal distributions in our probabilistic analyses (Table 2).

#### 2.7.2 Medication Error Costs and Quality-Adjusted Life Year (QALY) Decrements

Recent reviews found very few published estimates of the UK cost of a medication error [15, 16]. The cost estimates synthesised by one of these reviews [13] were not suitable for the present study, because they do not describe costs for different levels of pADE harm. Therefore, they would not allow use of the complete medication error frequency data available from our clinical study, which give an estimate of pADE harm that is specific to the context of the present investigation. Of the available estimates, the most appropriate for this study were those published by Karnon et al. [19], as they specifically relate to errors involving medicines administration in hospital, without restriction to specific types of medicines. In addition, Karnon’s data describe costs for three levels of pADE harm, as required for our model (Fig. 1). They include the costs of correcting an error detected before administration and treating the adverse effects of a pADE, including extra tests, treatments and hospital length of stay. Karnon’s data also include hypothetical QALY decrement ranges following a pADE based on limited data and research team discussion [21], which is fully described in their final report [32]. As this approach was not based on primary data, its validity is limited, but these data are the only available estimates. The Online Supplemental Material (page 21) describes how we used these data (uprated to 2018 values using the Hospital and Community Health Services Index [28]) to derive the model inputs for medication error costs and QALY decrements (Table 2). Given high levels of uncertainty about these parameters, additional sensitivity analyses considered scenarios where they were both higher and lower (Tables P and Q of the Online Supplemental Material). We assumed that other costs associated with preparing and administering a dose of voriconazole were identical to both guidelines and did not include them.

#### 2.7.3 User-Testing Costs

We calculated the NHS cost of user-testing using resource use data from our previous study (Table 2) [11]. This assumed an approach to user-testing similar to that currently used to write the IMG: user-testing shared between organisations within our perspective and performed by a hospital pharmacist with hospital nurses as interview participants. We assumed that other costs associated with preparing and delivering the IMG were identical to both guidelines and did not include them.

### 2.8 Model Analysis and Validation

We developed the model in Excel, using Monte Carlo simulation with 20,000 samples to propagate the joint uncertainty in model inputs into the outcomes, as this was found sufficient for graphical convergence of medication errors, costs and NMB. We reported costs in pounds (£) for the 2018 price year.

We established face validity by discussion with an advisory group (doctors, nurses, pharmacists and IMG advisory board members) and within the research team (health-economic modellers, medication-safety researchers and pharmacists). We systematically checked internal validity by assessing core calculations and applying a series of logical checks to ensure the direction of predictions was consistent. External and cross validity were more difficult to establish, as there are few comparable studies.
3 Results

The results from the base-case analysis for the whole cohort of patients within an NHS payer perspective are presented in Table 4. There were 157 fewer moderate-severe pADEs when user-tested guidelines were used, but with considerable uncertainty about this estimate. There was greater certainty around the reduction of 411 in the total number of pADEs when user-tested guidelines were used. There were fewer QALY decrements associated with the use of the user-tested guidelines and ICS favoured user-testing, with a 99% chance of being cost-saving. The cost-effectiveness plane (Fig. 2) indicates that 94.7% of simulations were clustered in the quadrant where user-testing was dominant. At a willingness-to-pay threshold of £20,000 per QALY, the NMB was over £3 million but with considerable uncertainty and a 96% chance of user-testing being cost-effective. The cost-effectiveness acceptability curve shows that this high likelihood of cost-effectiveness is maintained over a wide range of willingness-to-pay thresholds (Fig. 3).

Table 5 shows the results of selected sensitivity analyses (shown in full in Table M of the Online Supplemental Material). Uncertainty about the time horizon (Table 5), probability of detecting an error, probability of an error causing no harm, annual number of doses of intravenous voriconazole (Table 5), QALY decrement distributions and medication error cost distributions changes the magnitude of the outcomes but does not alter optimal strategy. Likewise, the magnitude of outcomes changes but the optimal strategy does not alter when all errors (rather than just the most severe error for each dose) are considered, or when the original NMB calculations used by Karnon et al. are applied. However, when the frequency of medication errors with current guidelines was reduced from 62% of doses (base case) to 32% of doses (similar to the frequency described by systematic reviews [3, 4]), the probability that user-testing is cost-effective is reduced and there is greater uncertainty about the ICS and NMB (Table 5). This reduction is even greater if the effectiveness of user-testing in reducing medication errors is also halved. The cost-effectiveness acceptability curves for these last two sensitivity analyses are shown in Fig. 3.

4 Discussion

This paper reports the development of a model to assess the cost-effectiveness of user-testing injectable medicines guidelines for use by nurses preparing intravenous medicines. The results suggest that for high-risk situations (e.g. an unfamiliar drug), user-testing reduces the number of pADEs and is both cost-saving and cost-effective. These findings were robust to many sources of uncertainty. For lower-risk situations, there is greater uncertainty about these outcomes, especially if user-testing is relatively less effective, although there is still > 70% probability that user-testing is cost-effective. Although these findings are specific to the UK, they are likely to be applicable in any setting that uses similar intravenous medicines administration systems, as user-testing is a relatively inexpensive intervention and a reduction in pADEs is likely to result in large cost reductions.

There are no other published economic evaluations of user-testing for improving medicines safety. The cost-effectiveness of several other hospital medicines safety interventions has been assessed. We based the model described in this study on one developed by Karnon et al., who used this

| Table 4 | Base-case analysis costs and outcomes of using current and user-tested guidelines to support the administration of intravenous voriconazole for the entire cohort of patients over a time horizon of 5 years |
|---------|---------------------------------------------------------------------------------------------------------------|
| Number of moderate-severe pADEs       | 885 (542–1330) | 728 (441–1102) | 157 (−13 to 363) | 0.96 |
| Total number of pADEs                 | 1202 (748–1775) | 792 (481–1102) | 411 (210 to 675) | 1.00 |
| QALY decrements                       | 238.8 (83.9–513.5) | 91.4 (26.0–237.6) | 147.5 (−24.9 to 406.1) | 0.95 |
| Total user-testing cost, £             | – | 6317 (6012–6627) | – | – |
| Health system costs, £                 | 1,225,800 (799,808–1,779,547) | 978,540 (635,010–1,434,142) | Incremental cost-saving: 240,943 (43,527 to 491,576) | 0.99 |
| Net monetary benefit, £             | – | – | 3,190,064 (−346,709 to 8,480,665) | 0.96 |

pADE preventable adverse drug effect, QALY quality-adjusted life year

aSuperiority corresponds to lower pADEs, costs, or QALY decrements

bWillingness-to-pay threshold of £20,000 per QALY

△ Adis
Table 5  Selected sensitivity analysis costs and outcomes of using current and user-tested guidelines to support the administration of intravenous voriconazole for the entire cohort of patients over a time horizon of 5 years. Table M of the Online Supplemental Material shows the results of all sensitivity analyses.

| Parameter varied from base-case analysis<sup>a</sup> | Mean decrease with user-testing (95% credible interval) | Incremental cost-saving | Net monetary benefit<sup>b</sup> |
|--------------------------------------------------|--------------------------------------------------------|-------------------------|---------------------------------|
|                                                  | Moderate-severe pADEs | All pADEs | QALY decrements | Mean (95% credible interval), £ | Probability user-testing cost-saving | Mean (95% credible interval), £ | Probability user-testing cost-effective |
| Base case (for comparison)                       | 157 (−13 to 363) | 411 (210–675) | 147.5 (−24.9 to 406.1) | 240,943 (43,527–491,576) | 0.99 | 3,190,064 (−346,709 to 8,480,665) | 0.96 |
| Time horizon reduced to 1 year                  | 31 (−3 to 73) | 82 (42–136) | 31.2 (−5.2 to 87.1) | 46,648 (3,177–101,006) | 0.98 | 669,944 (−74,287 to 1,815,983) | 0.96 |
| Time horizon increased to 10 years              | 315 (−32 to 728) | 823 (420–1,356) | 265.8 (−46.5 to 742.3) | 449,553 (79,168–911,049) | 0.99 | 5,766,532 (−612,727 to 15,470,264) | 0.96 |
| Number of doses of intravenous voriconazole administered using IMG per annum increased to 20,000 | 787 (−80 to 1,822) | 2,057 (1,039–3,402) | 730.5 (−133.4 to 2,041.3) | 1,231,112 (224,306–2,508,969) | 0.99 | 15,840,983 (−1,816,590 to 42,701,213) | 0.96 |
| Undetected error frequency with current guidelines reduced to 32%, with unchanged relative reduction in error frequency after user-testing | 80 (−77 to 252) | 208 (36–421) | 75.4 (−64.7 to 269.0) | 127,725 (−57,148–335,690) | 0.92 | 1,636,052 (−1,253,565 to 5,611,961) | 0.88 |
| Undetected error frequency with current guidelines reduced to 32% and relative effect of user-testing halved | 41 (−110 to 223) | 119 (−49–316) | 40.6 (−125.9 to 237.5) | 81,539 (−106,273–285,048) | 0.81 | 892,882 (−2,526,604 to 4,916,336) | 0.71 |

<sup>a</sup>See Tables I to L of the Online Supplemental Material

<sup>b</sup>Willingness-to-pay threshold = £20,000 per QALY

pADE preventable adverse drug event, QALY quality-adjusted life year
approach to evaluate the cost-effectiveness of three medicines safety interventions: computerised physician order entry (CPOE), additional ward pharmacists, and bar coding of medicines [21]. In contrast to the present study, it was uncertain whether any of these interventions were cost-saving, but all three were strongly cost-effective in terms of...
NMB. Karnon et al. later adapted this model to examine the cost-effectiveness of five medicines reconciliation interventions during hospital admission [33]. As in the present study, three interventions were found to be cost-saving while also increasing QALYs. A study using a different model concluded that the use of clinical pharmacists to detect errors in prescriptions for antineoplastic drugs was cost-effective, with estimated annual savings of €249,844 for one hospital [18]. Unlike the present study, other economic evaluations of hospital medicines safety projects have only considered intervention costs and not the costs avoided by reducing pADEs. For example, Vermeulen et al. estimated an incremental cost-effectiveness ratio (ICER) of €322.70 per pADE avoided by a CPOE system [17]. Based on the mean user-testing cost and mean total number of pADEs avoided (Table 4), the equivalent ICER for this study is £15.41 per pADE avoided.

These comparisons emphasise that user-testing is a relatively low-cost patient safety intervention, with a small initial cost and little requirement for repetition or maintenance (unlike training or CPOE for example). The present study assumed that user-testing was carried out by NHS-employed pharmacists. However, utilising the expertise of user-testing specialists, with the input of specialists in information writing and design, may produce even more effective guidelines. The time to complete the process might also be reduced, but with some increased costs. In addition, the present study describes an ‘intensive’ user-testing process, where the guideline for one medicine was tested with 30 nurses. However, some of the findings from this process are applicable to the IMG guidelines for all other medicines. Such learning will reduce the overall amount of testing required to improve the guidelines for all 350 medicines in the IMG, thus further improving the cost-effectiveness of user-testing.

4.1 Limitations

One limitation of our study is that the modelled results draw on an in situ simulation rather than the investigation of actual patient care [12]. This may have changed participants’ practice. Our previous simulation involved an unfamiliar medicine, whereas in day-to-day practice nurses are familiar with most of the medicines they administer. While the administration of lower-risk medicines was explored in the sensitivity analysis, it is not certain if these analyses were reasonable and explored the full range of possibilities. In addition, the classification of the potential level of harm caused by each error was based on the subjective judgement of an expert panel, although there was high agreement among this panel (Cronbach’s alpha 0.93), and it correctly assigned the level of harm for 13 of 15 medication errors with a known outcome [12]. A further limitation is the limited amount of empirical data used by Karnon et al. to derive their QALY decrements following a pADE and the lack of subsequent validation of these parameters [21], although sensitivity analyses found that the cost-effectiveness of user-testing was robust to large changes in these QALY decrements. Finally, a broader societal perspective would have included a greater range of costs, such as social care following permanent injury by a medication error.

4.2 Implications

These results suggest that the IMG should adopt the process of user-testing to produce guidelines for high-risk medicines. Evaluation of the impact of this change could inform a similar decision in relation to lower-risk medicines. As user-testing is a low-cost intervention and the costs of pADEs are relatively much higher, this approach should also be considered for other types of written guidelines for health professionals, including those both related and unrelated to medicines. Economic evaluation is an important tool to support the introduction of medicines safety interventions but is currently limited by lack of consensus and rigour in relation to how the costs and consequences of medication errors should be calculated [34]. Future research should aim to determine the most appropriate methods for calculating medication error costs and consequences in different contexts.

5 Conclusion

There is a high probability that user-testing injectable medicines guidelines for nurses preparing intravenous medicines in clinical areas is both cost-effective (96%) and cost-saving (99%) for high-risk medicines. The probability that user-testing is cost-effective and cost-saving is reduced for lower-risk medicines, but is still 71% and 80% (respectively), as user-testing is a low-cost intervention.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s40258-021-00675-z.

Acknowledgements Matthew Jones is funded by a National Institute for Health Research (NIHR) Transitional Research Fellowship for this research project. Howard Thom was supported by the NIHR Biomedical Research Centre at University Hospitals Bristol and Weston NHS Foundation Trust and the University of Bristol. Margaret Watson was funded by a Health Foundation Improvement Science Fellowship. This paper presents independent research funded by the National Institute for Health Research (NIHR), the NIHR Imperial Patient Safety Translational Research Centre and the NIHR Health Protection Research Unit in Healthcare Associated Infections and Antimicrobial Resistance at Imperial College in partnership with Public Health England (PHE). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, PHE or the Department of Health and Social Care. We are grateful to John Karnon (University of Adelaide) for providing copies of his original model and Robin Burfield (NHS Wales Informatics Service) for providing information on the usage of
the Injectable Medicines Guide. We also thank all the NHS organisations who provided information in response to our Freedom of Information request.

Declarations

Funding Matthew Jones is funded by a National Institute for Health Research (NIHR), Transitional Research Fellowship for this research project (grant number TRF-2017-10-006). Howard Thom was supported by the NIHR Biomedical Research Centre at University Hospitals Bristol and Weston NHS Foundation Trust and the University of Bristol. Margaret Watson was funded by a Health Foundation Improvement Science Fellowship. Bryony Dean Franklin is supported by the NIHR Imperial Patient Safety Translational Research Centre and the NIHR Health Protection Research Unit in Healthcare Associated Infections and Antimicrobial Resistance at Imperial College in partnership with Public Health England (PHE).

Conflict of Interests MDJ has received an honorarium from B. Braun Medical Ltd for speaking at an educational symposium on injectable medicines safety. DFK is academic advisor to Luto Research which develops, refines and tests health information materials. BDF has received honoraria from Pfizer for organising and chairing two educational symposia on medication safety, and recently supervised a Ph.D. student who was part funded by Cerner, an electronic health record systems vendor. HT has received consulting fees from Pfizer, Roche, Janssen, Novartis, and Bayer. For the remaining authors, none were declared.

Availability of Data and Material The model and supporting data created during this research are openly available from the University of Bath Research Data Archive at https://doi.org/10.15125/BATH-00960.

Code Availability The model created during this research is openly available from the University of Bath Research Data Archive at https://doi.org/10.15125/BATH-00960.

Ethics Approval and Consent to Participate The clinical study that provided many of the input data was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the University of Bath Research Ethics Approval Committee for Health (EPI17/18-192) and the United Kingdom’s Health Research Authority (IRAS 242505). Participants gave written, informed consent before taking part. This included consent to the publication of anonymised research data.

Author Contributions All authors contributed to the study conception and design. Primary data collection and identification of data inputs were performed by Matthew Jones. Model development and analysis were performed by Matthew Jones, Howard Thom and Rebecca Kandivaly. The first draft of the manuscript was written by Matthew Jones and all authors contributed to and approved the final manuscript.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Medication without harm - global patient safety challenge on medication safety. World Health Organization. 2017. https://apps.who.int/iris/rest/bitstreams/1083775/retrieve. Accessed 14 Nov 2019.
2. Lahue BJ, Pyenson B, Iwasaki K, Blumen HE, Forray S, Rothchild JM. National burden of preventable adverse drug events associated with inpatient injectable medications: healthcare and medical professional liability costs. Am Health Drug Benefits. 2012;5(7):1–10.
3. McLeod MC, Barber N, Franklin BD. Methodological variations and their effects on reported medication administration error rates. BMJ Qual Saf. 2013;22(4):278–89.
4. Keers RN, Williams SD, Cooke J, Ashcroft DM. Prevalence and nature of medication administration errors in health care settings: a systematic review of direct observational evidence. Ann Pharmacother. 2013;47(2):237–56.
5. Taxis K, Barber N. Ethnographic study of incidence and severity of intravenous drug errors. Br Med J. 2003;326(7391):684–7.
6. Wirtz V, Taxis K, Barber N. An observational study of intravenous medication errors in the United Kingdom and in Germany. Pharm World Sci. 2003;25:104–11.
7. Hartley GM, Dhillon S. An observational study of the prescribing and administration of intravenous drugs in a general hospital. Int J Pharm Pract. 1998;6(1):38–45.
8. Lyons I, Furniss D, Blandford A, Chumbley G, Iacovides I, Wei L, et al. Errors and discrepancies in the administration of intravenous infusions: a mixed methods multihospital observational study. BMJ Qual Saf. 2018;27:892–901.
9. Lawton R, McEachan RR, Giles SJ, Sirriyeh R, Watt IS, Wright J. Development of an evidence-based framework of factors contributing to patient safety incidents in hospital settings: a systematic review. BMJ Qual Saf. 2012;21(5):369–80.
10. Chang A, Schyve PM, Croteau RJ, O’Leary DS, Loeb JM. The JCAHO patient safety event taxonomy: a standardized terminology and classification schema for near misses and adverse events. Int J Qual Health Care. 2005;17(2):95–105.
11. Jones MD, Franklin BD, Watson MC, Raynor DK. User testing to improve retrieval and comprehension of information in guidelines to improve medicines safety. J Patient Saf. 2020. https://doi.org/10.1097/PTS.0000000000000723.
12. Jones MD, McGrogan A, Raynor DKT, Watson MW, Franklin BD. User-testing guidelines to improve the safety of intravenous medicines administration: a randomised in-situ simulation study. BMJ Qual Saf. 2020;30:17–26.
13. Elliott RA, Camacho E, Jankovic D, Sculpher MJ, Faria R. Economic analysis of the prevalence and clinical and economic burden of medication error in England. BMJ Qual Saf. 2021;30:96–105.
14. Peto T, Gray A. Economic evaluation using decision analytical modelling: design, conduct, analysis, and reporting. BMJ. 2011;342:d1766.
15. Walsh EK, Hansen CR, Sahm LJ, Kearney PM, Doherty E, Bradley CP. Economic impact of medication error: a systematic review. Pharmacoeconimol Drug Saf. 2017;26(5):481–97.
16. Elliott RA, Camacho E, Campbell F, Jankovic D, Martyn-St James M, Kalentehler E, et al. Prevalence and economic burden of medication errors in the NHS in England: Policy Research Unit in Economic Evaluation of Health and Care Interventions.
17. Vermeulen KM, van Doormaal JE, Zaal RJ, Mol PG, Lenderink AW, Haaijer-Ruskamp FM, et al. Cost-effectiveness of an electronic medication ordering system (CPOE/CDSS) in hospitalized patients. Int J Med Inform. 2014;83(8):572–80.

18. Nerich V, Borg C, Villanueva C, Thiery-Vuillemin A, Helias P, Rohrlch PS, et al. Economic impact of prescribing error prevention with computerized physician order entry of injectable antineoplastic drugs. J Oncol Pharm Pract. 2013;19(1):8–17.

19. Karnon J, McIntosh A, Dean J, Bath P, Hutchinson A, Oakley J, et al. A prospective hazard and improvement analytic approach to predicting the effectiveness of medication error interventions. Saf Sci. 2007;45(4):523–39.

20. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. BMJ. 2013;346:f1049.

21. Karnon J, McIntosh A, Dean J, Bath P, Hutchinson A, Oakley J, et al. Modelling the expected net benefits of interventions to reduce the burden of medication errors. J Health Serv Res Policy. 2008;13(2):85–91.

22. Samp JC, Touchette DR, Marinac JS, Kuo GM, American College of Clinical Pharmacy Practice-Based Research Network C. Economic evaluation of the impact of medication errors reported by U.S. clinical pharmacists. Pharmacotherapy. 2014;34(4):350–7.

23. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013. 2013. https://www.nice.org.uk/process/PBG9/chapter/Foreword. Accessed 17 Oct 2019.

24. Douglass AM, Elder J, Watson R, Kallay T, Kirsh D, Robb WG, et al. A randomized controlled trial on the effect of a double check on the detection of medication errors. Ann Emerg Med. 2018;71(1):74–82 e1.

25. Modic MB, Albert NM, Sun Z. Does an insulin double-checking procedure improve patient safety? J Nurs Adm. 2016;46(3):154–60.