
Article:
Durojaiye, O.C., Morgan, R., Chelaghma, N. et al. (5 more authors) (2021) External validity and clinical usefulness of a risk prediction model for 30 day unplanned hospitalization in patients receiving outpatient parenteral antimicrobial therapy. Journal of Antimicrobial Chemotherapy, 76 (8). pp. 2204-2212. ISSN 0305-7453

https://doi.org/10.1093/jac/dkab127

JAC-2021-0059-R1-FINAL https://doi.org/10.1093/jac/dkab127

External validity and clinical usefulness of a risk prediction model for 30 day unplanned hospitalization in patients receiving outpatient parenteral antimicrobial therapy

Running title: Validation of risk prediction model for OPAT readmission

Oyewole Christopher DUROJAIYE [1,2]*, Robin MORGAN[1], Naziha CHELAGHMA[3], Joyeeta PALIT[1], Christopher KEIL[4], Rasha OMER[1], Katharine CARTWRIGHT[1] and Evangelos I. KRITSOTAKIS[5,6]

[1]*Department of Infection and Tropical Medicine, Royal Hallamshire Hospital, Sheffield, UK;*

[2]*Department of Microbiology, Royal Derby Hospital, Derby, UK;*

[3]*Department of Cardiology, University Hospitals of Derby and Burton NHS Foundation Trust, Burton-on-Trent, UK;*

[4]*Sheffield Teaching Hospital NHS Foundation Trust, Sheffield, UK;*

[5]*Laboratory of Biostatistics, School of Medicine, University of Crete, Heraklion, Greece;*

[6]*School of Health and Related Research, Faculty of Medicine, Dentistry and Health, The University of Sheffield, Sheffield, UK*

*Corresponding author. E-mail: docwolex@yahoo.co.uk

Objectives: Outpatient parenteral antimicrobial therapy (OPAT) is increasingly used to treat a variety of infections. However, hospital readmissions remain relatively common. We examined the external validity and clinical usefulness of a previously derived risk prediction model for 30 day unplanned hospitalization in patients receiving OPAT.

Methods: A retrospective cohort study was conducted at two large teaching hospitals in the UK. The design comprised quasi-external temporal validation on patients from the same OPAT setting as the model development, and broader external validation on patients from a different setting. The model predictors were age, prior hospitalizations in the preceding 12 months, Charlson comorbidity score, concurrent IV antimicrobial therapy, type of infection and mode of OPAT treatment. Discriminative ability, calibration and clinical usefulness were assessed.

Results: Data from 2578 OPAT patients were analysed. The rates of 30 day unplanned hospitalization were 11.5% (123/1073), 12.9% (140/1087) and 25.4% (106/418) in the model derivation, temporal validation and broader external validation cohorts, respectively. The discriminative ability of the prediction model was adequate on temporal validation (*c*-statistic 0.75; 95% CI: 0.71–0.79) and acceptable on broader validation (*c*-statistic 0.67; 95% CI: 0.61–0.73). In both external cohorts, the model displayed excellent calibration between observed and predicted probabilities. Decision curve analysis showed increased net benefit across a range of meaningful risk thresholds.

Conclusions: A simple risk prediction model for unplanned readmission in OPAT patients demonstrated reproducible predictive performance, broad clinical transportability and clinical usefulness. This model may help improve OPAT outcomes through better identification of high-risk patients and provision of tailored care.

Introduction

IV antimicrobial agents are increasingly administered in outpatient settings to treat a variety of infections in patients who need parenteral therapy but are well enough not to require hospitalization.[1,2] The efficacy and safety of outpatient parenteral antimicrobial therapy (OPAT) have been well documented.[3–6] Despite its benefits, OPAT is potentially associated with increased clinical risk due to

the reduced level of clinical supervision and monitoring compared with inpatient care. Even with careful patient selection and multidisciplinary team-driven therapeutic plans, the nature of infections treated, the use of potentially toxic antimicrobial agents and the duration of treatment imply that complications, including readmissions for some patients, are inevitable. Thirty day readmission rates have been used in the UK and internationally as a marker of healthcare quality and OPAT outcome.[7–9] Predicting and preventing unplanned hospitalization could improve patient outcomes and reduce healthcare costs. We previously developed a risk prediction model for 30 day unplanned hospitalization in patients receiving OPAT at a large teaching hospital in Sheffield, UK.[10] The model predictors were: age; Charlson comorbidity score; prior hospitalizations in the preceding 12 months; concurrent IV antimicrobial therapy; type of infection; and mode of OPAT treatment. The performance of clinical prediction models can vary due to changing outcome rates, shifting patient mix and evolving clinical practice. Before a prediction model can be used in clinical practice, it is essential to evaluate its ability to produce accurate predictions as well as its usefulness to support clinical decision-making on new subjects from the same source population (reproducibility) and in different clinical settings (transportability).[11–13]

This study aimed to evaluate the external validity and clinical usefulness of a previously published model for predicting the risk of 30 day unplanned hospitalization in patients receiving OPAT,[10] by selecting more recently treated subjects in the OPAT setting where the model was derived (reproducibility) and subjects from a different OPAT setting (transportability).

Patients and methods

Study design and setting

We conducted a retrospective cohort study of adult patients (aged >18 years) who received OPAT at two large teaching hospitals in the UK (Sheffield Teaching Hospitals and Royal Derby Hospital). The design comprised quasi-external validation on patients (Sheffield cohort) having temporal characteristics different from those of the model development (temporal validation) and broader

external validation on patients from a different OPAT setting (Derby cohort) with expected differences in case mix.

The Sheffield and Derby OPAT services were formally established in 2006 and 2013, respectively. Both services are run by a multidisciplinary team of infection specialists, specialist nurses and clinical antimicrobial pharmacists. In the Sheffield OPAT centre, antimicrobials were delivered by three distinct pathways: daily attendance at the 'infusion centre'; self or carer administration in the patient's home; and administration by a district/community nurse in the patient's home. However, in the Derby OPAT setting, antimicrobials were primarily delivered in the patient's home by visiting nurses. Both OPAT centres maintain electronic databases to prospectively record patient demographics, clinical diagnosis, model of delivery, antimicrobial agents, treatment duration, type of vascular access, clinical outcome and complications. Patient selection, antimicrobial regimen and mode of OPAT delivery were the responsibility of the infection specialists at each centre. The clinical responsibility for patients receiving OPAT and their follow-up were shared between the referring clinicians and the OPAT infection specialists, unless otherwise agreed.

The temporal and broader external validation cohorts were made up of data extracted from the OPAT databases and electronic health records of patients who received OPAT between January 2018 and January 2020. Age (years) was determined at the time of commencing OPAT. Weighted Charlson comorbidity score was calculated for each patient and was determined at the time OPAT was commenced.[14]

Original risk prediction model

The original prediction model was developed using data from 1073 patients who received OPAT between January 2015 and January 2017 in Sheffield (derivation cohort).[10] The primary outcome was 30 day unplanned hospitalization, defined as unplanned inpatient admission to an acute care hospital for any reason within 30 days of discharge from the OPAT service. The model consisted of six

predictors: age; Charlson comorbidity score;[14] number of prior non-OPAT hospitalizations in the preceding 12 months; concurrent receipt of more than one IV antimicrobial agent; type of infection; and mode of OPAT delivery (infusion centre, community nurse or self/carer administration). The linear predictor (LP) for a patient was given by: LP=−3.628+(0.016×age in years)+(0.264×number of prior hospitalizations)+(0.103×Charlson comorbidity score)+(0.248, if self/carer administration)+(0.479, if infusion centre)+(0.635, if IV combination therapy)+(0.480, if endovascular infection)−(0.337, if respiratory disease)+(0.189, if urogenital infection)−(0.037, if bone and joint infection)−(0.776, if skin and soft tissue infection). The probability (or risk) of 30 day unplanned hospitalization for the same patient was given by: 1/[1+exp(−LP)].

Sample size

For external validation of prognostic models, a minimum of 100 outcome events is recommended to ensure adequate power to detect changes in predictive performance in external datasets.[15] Given our previous experience with the Sheffield OPAT service (about 65 unplanned readmissions per year),[10] we reviewed 2 years' worth of OPAT records to form the external validation cohorts in this study.

Statistical analysis

We conducted external validation in line with the methodological framework proposed by Debray *et al*.[13] We report the study in accordance with the transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) recommendations.[12] Because some patients had more than one episode of OPAT treatment during the study period, we performed individual-level analysis taking a simple random sample of one OPAT episode per patient.

The first step in our validation approach was to assess the extent to which the external cohorts are useful for evaluating the model's reproducibility or its transportability.[13] The relatedness between the derivation cohort and the external cohorts was reviewed using two approaches. First, the distribution

of context-important patient characteristics, including model predictors and outcomes, were compared. Second, the degree of relatedness in case mix between the development and validation cohorts was quantified by the *c*-statistic [area under the receiver operating characteristic (ROC) curve] of a binary logistic regression (membership) model predicting the probability that an individual patient is a member of the derivation cohort as opposed to the validation cohort.[13] Independent variables in the membership model were the outcome and the six predictors of the original predictive model for unplanned hospitalization. Low values of the concordance *c*-statistic for the membership model (close to 0.5) indicate indistinguishable case mix between the datasets.

In the second step, we examined heterogeneity in predictor–outcome associations by fitting the original set of predictors internally to each validation cohort and by calculating optimism-corrected performance measures using bootstrap resampling with 500 replications.[16] Optimism-corrected measures were reduced by the estimated deterioration that the model is expected to have when applied to new individuals. We then externally assessed the performance of the originally developed model in each validation cohort. For the latter, a logistic regression model with the unplanned hospitalization as the outcome variable and the linear predictor of the original risk model as the only covariate was applied to each validation dataset.

Model performance was evaluated in terms of discriminative ability (to differentiate patients with unplanned hospitalization from those without) and calibration (agreement between predicted probabilities and actual rates of unplanned hospitalization). Discrimination was quantified with the *c*-statistic and was displayed graphically with an ROC curve. Calibration was broadly assessed by the Hosmer–Lemeshow goodness-of-fit test; a non-significant test indicates good calibration. The deviation of the intercept (calibration-in-the-large) and the slope of the calibration line (plotting predicted against observed event rates) from the ideal values of 1 and 0, respectively, were examined.[17] Calibration-in-the-large represents the agreement between the overall observed and the overall predicted risk of hospitalization (by definition this is zero in the development sample). Calibration slopes >1 typically occur when predicted probabilities do not vary enough (e.g. predicted

risks are systematically too low), whereas slopes <1 occur when they vary too much (e.g. low predictions are too low and high predictions are too high).[13,17] Calibration was visually inspected using loess-based calibration plots.[18]

In the third step, we used recalibration methods to improve the performance of the originally developed model in the validation cohorts.[19] Where calibration-in-the-large was significantly different from zero due to different overall risks of hospitalization in the external cohorts, intercept recalibration was performed by fitting a new logistic regression model with an intercept only and an offset term for the linear predictor of the original risk model.

Finally, we performed decision curve analysis to assess the clinical usefulness of the recalibrated model on the external cohorts.[20] This analysis provides insight into the range of predicted risks (decision thresholds) for which the model has a greater net benefit (NB) than an 'intervention for all' strategy (assuming all patients have complications and require hospitalization) or an 'intervention for none' strategy (assuming none have complications). NB combines the benefits of true positives and the harms of false positives on a single scale by weighting false positives by the odds of the chosen risk threshold to select patients for intervention.[20]

All patient data were anonymized prior to analysis. Data were processed and analysed using Stata/MP 14.1 (StataCorp, College Station, TX, USA). The study was approved by the local clinical effectiveness unit as part of ongoing commitment to service development.

Results

A total of 2578 patients were identified, of which 1073 were recorded in the initial model derivation cohort. We recorded 418 and 1087 patients in the temporal validation cohort (Sheffield, 2018–20) and broader external validation cohort (Derby, 2018–20), respectively. Table 1 compares patient

characteristics and outcomes between the development and validation stages. The temporal validation (Sheffield) cohort had similar patient characteristics and outcomes as the derivation cohort. Of note, more subjects in the temporal validation cohort received glycopeptides (14.8% versus 9.1%; $P<0.001$) and concurrent oral antibiotic therapy (20.1% versus 11.4%; $P<0.001$) but fewer received cephalosporins (66.0% versus 73.6%; $P<0.001$), prior OPAT treatment (7.5% versus 16.1%; $P<0.001$) and 'infusion centre' administration (66.0% versus 73.6%; $P<0.001$) than the model development subjects. Rates of 30 day unplanned hospitalization were similar in the two cohorts (11.5% versus 12.9%; $P=0.314$). The *c*-statistic for the membership model comparing the respective cohorts was 0.60 (95% CI: 0.57–0.62) indicating similar case mix. This implies that temporal validation in this study merely assesses the model's reproducibility in the same target population rather than transportability in a different setting.

However, the broader external validation (Derby) cohort was different to the model derivation cohort. Subjects in the Derby cohort were older (mean age 68 versus 56 years; $P<0.001$), with more comorbidities (median Charlson score 2 versus 1; $P<0.001$), MDR organisms (11.7% versus 8.0%; $P=0.026$), central vascular access (88.8% versus 70.6%; $P<0.001$), more severe indications for OPAT and with a higher rate of unplanned hospitalization (25.4% versus 11.5%; $P<0.001$). The *c*-statistic for the membership model was 0.95 (95% CI: 0.94–0.96) indicating highly discordant case mix. This implies that the broader external validation in this study reflects the transportability of the original risk prediction model to more severe patients selected differently to receive OPAT.

The effects of the six predictors in the original risk model were consistent (in direction and magnitude) across the three cohorts (Table 2). Internal validation of the six-predictor model, when separately fitted in each validation cohort, showed adequate discriminative performance with optimism-corrected *c*-statistics at 0.70, 0.78 and 0.74 in the derivation, temporal validation and broader validation cohorts, respectively. The Hosmer–Lemeshow goodness-of-fit test indicated good broad calibration in deciles of predicted risks in all cohorts. There were no significant differences in the calibration intercept and calibration slope from the ideal values of 0 and 1, respectively.

When the originally derived model was externally validated, adequate discriminative ability was retained in the temporal validation cohort (*c*-statistic 0.75; 95% CI: 0.71–0.79) and despite being reduced, discriminative ability was acceptable in the broader validation cohort (*c*-statistic 0.67; 95% CI: 0.61–0.73). There were no significant differences in the calibration slopes from the ideal value of 1 in either external cohort. However, calibration-in-the-large was slightly higher than zero in temporal validation and too high in broader external validation (intercept 0.54; 95% CI: 0.31–0.77; Hosmer–Lemeshow test $P<0.001$). ROC curves and calibration plots are contrasted in Figure 1. Poor calibration-in-the-large was easily overcome by recalibrating the intercept of the model. The intercept-recalibrated model demonstrated excellent calibration performance in the external cohorts (Figure 1).

Decision curve analysis showed that the recalibrated model has greater NB (i.e. clinically useful) than 'intervention for all' or 'intervention for none' strategies in a range of meaningful predicted probability thresholds between approximately 15% and 50% for both external validation cohorts (Figure 2).

Discussion

This study examined the temporal and broader external validity and clinical usefulness of a previously published risk prediction model for 30 day unplanned hospitalization in patients receiving OPAT. The rates of unplanned hospitalization in the model derivation cohort, temporal validation cohort and broader validation cohort (11%, 13% and 25%, respectively) are comparable with other OPAT studies.[8,9,21,22] The higher rate of unplanned hospitalization in the broader validation (Derby) cohort is likely due to the severity of infections treated, older age and high levels of comorbidity among the subjects. The discriminatory performance of the prediction model was retained and good in the temporal validation cohort (*c*-statistic 0.75), meaning that all subjects can be discriminated

appropriately between the readmitted and non-readmitted group. The performance was reduced in the broader external dataset (*c*-statistic 0.67; 95% CI: 0.61–0.73), but remained potentially acceptable. In general, a *c*-statistic of 0.5 indicates no discrimination; 0.7 to 0.8 shows acceptable discrimination; 0.8 to 0.9 shows excellent discrimination; and values ≥0.9 are considered outstanding discrimination but are rarely seen in practice.[23] In both external validation cohorts, the calibration slopes were very close to the ideal value of 1, suggesting good correlation between observed and predicted risks across the subjects. However, calibration-in-the-large was poor in the broader validation cohort, indicating systematic overprediction of the risk of unplanned hospitalization in that cohort.

In external validation studies, decreases in performance of prediction models are common and can be due to differences in case mix of patients between the development and validation cohorts, overfitting of the model to the data used for development, differences in the effect of the model predictors between the development and validation cohorts, or a combination of these factors.[13,24] Prediction model performance can sometimes be improved by recalibration to the new setting, re-estimation of regression coefficients or including more predictors.[19,24] We addressed the poor calibration-in-the-large by recalibrating the intercept of the model to reflect the higher overall risks of unplanned hospitalization in the external cohorts (especially the broader validation cohort) compared with the derivation cohort.

Clinical usefulness is the ability to make a better decision when using a prediction model compared with when not using the model.[20,24] However, it is not clear how good the model's discriminative and calibration performances need to be to warrant clinical use. In this study, we confirmed the clinical usefulness of the recalibrated model by calculating its NB. NB is a simple type of decision analysis; similar to the idea of net profit in business.[20] We found the NB difference (true NB) of using the predictive model to be greater than 0 for all threshold probabilities between 0.15 and 0.50. Thus, our study demonstrates that the model is useful for identifying patients at high risk of unplanned hospital readmission upon presentation to an OPAT service. Based on our results, the model can be used to aid clinicians in managing patients receiving IV antimicrobial therapy in outpatient settings. Once

identified as ‘high risk’, careful forward planning can help to prevent hospital readmissions and optimize clinical outcomes—by careful patient selection, closer monitoring and timely follow-up.[25]

This study has limitations that should be acknowledged. Caution should be applied when using the model. Each OPAT centre may have a different case mix of patients, OPAT structure and mode of delivery from those of the centres in which our study was conducted. Therefore, each OPAT service should compare their case mix and structure with those of the study OPAT services before using the risk prediction model directly. Recalibration of the model should be considered in settings with broader case mix before clinical use. Our analysis was retrospective, but the data were originally collected prospectively, which reduces the risk of measurement bias or poor accuracy of recorded data. However, we cannot be entirely confident that we have not missed some patients who were readmitted to other hospitals. Nevertheless, most patient interactions with healthcare systems are well documented in their clinical records. In the model development, we did not explore factors (such as therapeutic drug levels, frequency of monitoring or follow-up visits) that are not readily available prior to OPAT but are plausible risk factors for readmission. Our aim was to develop a risk prediction model based on parameters available on presentation to an OPAT service.

In conclusion, the prediction model is temporally and externally valid, and clinically useful for the prediction of 30 day unplanned hospitalization in patients receiving OPAT. It may help improve OPAT outcomes through better identification of high-risk patients and provision of tailored care. Future research should focus on recalibration and assessing the model performance in other OPAT centres, especially those with broader case mix from the study cohorts.

Acknowledgements

We wish to thank all the clinical and administrative staff working in the Sheffield and Derby OPAT service.

Funding

This study was carried out as part of our routine work.

Transparency declarations

None to declare.

Author contributions

O.C.D. and E.I.K. conceptualized and designed the study. K.C. contributed to the conduct of the study. O.C.D., R.M., N.C., J.P., C.K. and R.O. collected and extracted data. E.I.K. performed the statistical analysis. O.C.D. and E.I.K. drafted the manuscript. All authors contributed to the final manuscript.

References

1 Chapman ALN, Patel S, Horner C *et al*. Updated good practice recommendations for outpatient parenteral antimicrobial therapy (OPAT) in adults and children in the UK. *JAC-Antimicrob Resist* 2019; **1**: dlz026.

2 Durojaiye OC, Cartwright K, Ntziora F. Outpatient parenteral antimicrobial therapy (OPAT) in the UK: a cross-sectional survey of acute hospital trusts and health boards. *Diagn Microbiol Infect Dis* 2019; **93**: 58–62.

3 Durojaiye OC, Bell H, Andrews D *et al*. Clinical efficacy, cost analysis and patient acceptability of outpatient parenteral antibiotic therapy (OPAT): a decade of Sheffield (UK) OPAT service. *Int J Antimicrob Agents* 2018; **51**: 26–32.

4 Barr DA, Semple L, Seaton RA. Outpatient parenteral antimicrobial therapy (OPAT) in a teaching hospital-based practice: a retrospective cohort study describing experience and evolution over 10 years. *Int J Antimicrob Agents* 2012; **39**: 407–13.

5 Quintens C, Steffens E, Jacobs K *et al*. Efficacy and safety of a Belgian tertiary care outpatient parenteral antimicrobial therapy (OPAT) program. *Infection* 2020; **48**: 357–66.

6 Shrestha NK, Blaskewicz C, Gordon SM *et al*. Safety of outpatient parenteral antimicrobial therapy in nonagenarians. *Open Forum Infect Dis* 2020; **7**: ofaa398.

7 Rumball-Smith J, Hider P. The validity of readmission rate as a marker of the quality of hospital care, and a recommendation for its definition. *N Z Med J* 2009; **122**: 63–70.

8 Allison GM, Muldoon EG, Kent DM *et al*. Prediction model for 30-day hospital readmissions among patients discharged receiving outpatient parenteral antibiotic therapy. *Clin Infect Dis* 2014; **58**: 812–9.

9 Huang V, Ruhe JJ, Lerner P *et al*. Risk factors for readmission in patients discharged with outpatient parenteral antimicrobial therapy: a retrospective cohort study. *BMC Pharmacol Toxicol* 2018; **19**: 50.

10 Durojaiye OC, Kritsotakis EI, Johnston P *et al*. Developing a risk prediction model for 30-day unplanned hospitalization in patients receiving outpatient parenteral antimicrobial therapy. *Clin Microbiol Infect* 2019; **25**: 905.e1–7.

11 Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. *Ann Intern Med* 1999; **130**: 515–24.

12 Moons KGM, Altman DG, Reitsma JB *et al*. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* 2015; **162**: W1–73.

13 Debray TPA, Vergouwe Y, Koffijberg H *et al*. A new framework to enhance the interpretation of external validation studies of clinical prediction models. *J Clin Epidemiol* 2015; **68**: 279–89.

14 Charlson ME, Pompei P, Ales KL *et al*. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987; **40**: 373–83.

15 Collins GS, Ogundimu EO, Altman DG. Sample size considerations for the external validation of a multivariable prognostic model: a resampling study. *Stat Med* 2016; **35**: 214–26.

16 Steyerberg EW, Bleeker SE, Moll HA *et al*. Internal and external validation of predictive models: a simulation study of bias and precision in small samples. *J Clin Epidemiol* 2003; **56**: 441–7.

17 Stevens RJ, Poppe KK. Validation of clinical prediction models: what does the "calibration slope" really measure? *J Clin Epidemiol* 2020; **118**: 93–9.

18 Austin PC, Steyerberg EW. Graphical assessment of internal and external calibration of logistic regression models by using loess smoothers. *Stat Med* 2014; **33**: 517–35.

19 Janssen KJM, Moons KGM, Kalkman CJ *et al*. Updating methods improved the performance of a clinical prediction model in new patients. *J Clin Epidemiol* 2008; **61**: 76–86.

20 Vickers AJ, Van Calster B, Steyerberg EW. Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. *BMJ* 2016; **352**: i6.

21 Means L, Bleasdale S, Sikka M *et al*. Predictors of hospital readmission in patients receiving outpatient parenteral antimicrobial therapy. *Pharmacotherapy* 2016; **36**: 934–9.

22 Bengoetxea I, Onaindia MJ, Apezetxea A *et al*. Predictive clinical rule for readmissions in OPAT. Improving in security. *Open Forum Infect Dis* 2017; **4**: S335.

23 Hosmer DW, Lemeshow S, Sturdivant RX. *Applied Logistic Regression*. 3rd ed. John Wiley & Sons Inc., 2013.

24 Steyerberg EW. *Clinical Prediction Models: a Practical Approach to Development, Validation and Updating.* 2nd ed. Springer, 2019.

25 Saini E, Ali M, Du P *et al*. Early ID outpatient follow-up of OPAT patients reduces 30-day readmission. *Open Forum Infect Dis* 2017; **4**: S332.

Table 1. Characteristics of subjects between development and validation cohorts

Characteristics	Derivation cohort (*n*=1073)	Temporal validation cohort (*n*=1087)	*P* value[a]	Broader validation cohort (*n*=418)	*P* value[b]
Male sex, *n* (%)	611 (56.9)	641 (59.0)	0.340	259 (62.0)	0.078
Age, years, mean (SD)	56 (17.5)	56 (17.7)	0.795	68 (14.4)	<0.001
Age group, *n* (%)			0.702		<0.001
≤30	97 (9.0)	112 (10.3)		9 (2.2)	
31–40	121 (11.3)	108 (9.9)		15 (3.6)	
41–50	177 (16.5)	164 (15.1)		22 (5.3)	
51–60	209 (19.5)	225 (20.7)		58 (13.9)	
61–70	219 (20.4)	228 (21.0)		116 (27.8)	
>70	250 (23.3)	250 (23.0)		198 (47.4)	
Comorbidities, *n* (%)					
Chronic pulmonary disease	197 (18.4)	111 (10.2)	<0.001	94 (22.5)	0.071
Diabetes with complications	109 (10.2)	123 (11.3)	0.385	100 (23.9)	<0.001
Peripheral vascular disease	87 (8.1)	59 (5.4)	0.014	62 (14.8)	<0.001
Diabetes without complications	107 (10.0)	97 (8.9)	0.405	62 (14.8)	0.008
Tumour without metastasis	90 (8.4)	62 (5.7)	0.015	25 (6.0)	0.119
Moderate or severe renal disease	82 (7.6)	134 (12.3)	<0.001	80 (19.1)	<0.001
Connective tissue disease	71 (6.6)	56 (5.2)	0.149	29 (6.9)	0.824
Myocardial infarction	79 (7.4)	68 (6.3)	0.308	45 (10.8)	0.034
Cerebrovascular disease	49 (4.6)	64 (5.9)	0.169	31 (7.4)	0.030
Congestive heart failure	53 (4.9)	62 (5.7)	0.429	32 (7.7)	0.044
Peptic ulcer disease	40 (3.7)	10 (0.9)	<0.001	7 (1.7)	0.047
Moderate or severe liver disease	25 (2.3)	24 (2.2)	0.849	11 (2.6)	0.733
Metastatic solid tumour	21 (2.0)	20 (1.8)	0.842	12 (2.9)	0.284
Lymphoma	8 (0.7)	6 (0.6)	0.576	4 (1.0)	0.682
Leukaemia	11 (1.0)	12 (1.1)	0.858	6 (1.4)	0.505
Hemiplegia	6 (0.6)	8 (0.7)	0.610	1 (0.2)	0.431
Dementia	6 (0.6)	9 (0.8)	0.455	8 (1.9)	0.022
AIDS	2 (0.2)	0 (0.0)	0.999	0 (0.0)	0.999
Mild liver disease	0 (0.0)	27 (2.5)	0.998	4 (1.0)	0.999
Charlson comorbidity score, median (IQR)	1 (0–2)	0 (0–2)	0.308	2 (1–3)	<0.001
Charlson comorbidity score, *n* (%)			<0.001		<0.001
0	472 (44.0)	564 (51.9)		104 (24.9)	
1	203 (18.9)	154 (14.2)		83 (19.9)	

2	170 (15.8)	156 (14.4)		83 (19.9)	
3	101 (9.4)	64 (5.9)		49 (11.7)	
4+	127 (11.8)	149 (13.7)		99 (23.7)	
Indication for OPAT, *n* (%)			0.267		0.000
Skin and soft tissue infection	616 (57.4)	584 (53.7)		4 (1.0)	
Bone and joint infection	137 (12.8)	174 (16.0)		242 (57.9)	
Urogenital infection	70 (6.5)	63 (5.8)		6 (1.4)	
Respiratory disease	45 (4.2)	46 (4.2)		68 (16.3)	
Endovascular infection	45 (4.2)	43 (4.0)		40 (9.6)	
Other indication	160 (14.9)	177 (16.3)		58 (13.9)	
MDR organism, *n* (%)	86 (8.0)	89 (8.2)	0.883	49 (11.7)	0.026
Mode of antimicrobial delivery, *n* (%)			<0.001		<0.001
Infusion centre	767 (71.5)	673 (61.9)		3 (0.7)	
Community nurse[c]	201 (18.7)	289 (26.6)		412 (98.6)	
Self/carer administration	105 (9.8)	125 (11.5)		3 (0.7)	
Type of vascular access, *n* (%)			0.267		<0.001
Central line	758 (70.6)	744 (68.4)		371 (88.8)	
Peripheral access	315 (29.4)	343 (31.6)		47 (11.2)	
Antimicrobial agent, *n* (%)[d]					
Penicillin	93 (8.7)	99 (9.1)	0.719	121 (28.9)	<0.001
Cephalosporin	790 (73.6)	717 (66.0)	<0.001	111 (26.6)	<0.001
Carbapenem	104 (9.7)	121 (11.1)	0.274	59 (14.1)	0.015
Glycopeptide	98 (9.1)	161 (14.8)	<0.001	134 (32.1)	<0.001
Other	51 (4.8)	68 (6.3)	0.127	15 (3.6)	0.328
Concurrent IV OPAT, *n* (%)	81 (7.5)	81 (7.5)	0.932	21 (5.0)	0.085
Oral antibiotic included, *n* (%)	122 (11.4)	218 (20.1)	<0.001	111 (26.6)	<0.001
Duration of OPAT, days, median (IQR)	7 (4–19)	7 (3–21)	0.776	18 (8–32)	<0.001
Number of prior hospitalizations, median (IQR)[e]	0 (0–1)	0 (0–1)	0.096	1 (0–2)	0.011
Prior OPAT stay in past 12 months, *n* (%)	173 (16.1)	81 (7.5)	<0.001	42 (10.0)	0.003
Outcomes					
30 day hospitalization, *n* (%)	145 (13.5)	159 (14.6)	0.457	117 (28.0)	<0.001
30 day unplanned hospitalization, *n* (%)	123 (11.5)	140 (12.9)	0.314	106 (25.4)	<0.001
Reason for 30 day unplanned admission, *n* (%)			0.261		0.042
Worsening of infection/no improvement	52 (42.3)	64 (45.7)		35 (33.0)	
Non-OPAT related	50 (40.7)	40 (28.6)		36 (34.0)	
New infection	8 (6.5)	14 (10.0)		18 (17.0)	
Adverse drug reaction	7 (5.7)	14 (10.0)		12 (11.3)	
IV line-related complications	3 (2.4)	6 (4.3)		5 (4.7)	

Clostridioides difficile-associated diarrhoea	2 (1.6)	2 (1.4)	0 (0.0)
Unknown	1 (0.8)	0 (0.0)	0 (0.0)

[a]*P* value refers to the comparison between the Sheffield derivation cohort and the Sheffield validation cohort.

[b]*P* value refers to the comparison between the Sheffield derivation cohort and the Derby validation cohort.

[c]Community nurse refers to administration of antimicrobial therapy in a patient's home by a nurse.

[d]Some patients received more than one antimicrobial agent. Thus, total number of antimicrobial agents is greater than total number of patients.

[e]In 12 months preceding current OPAT episode.

Table 2. Multivariable logistic regression analysis for the risk of 30 day unplanned hospitalization in patients receiving OPAT

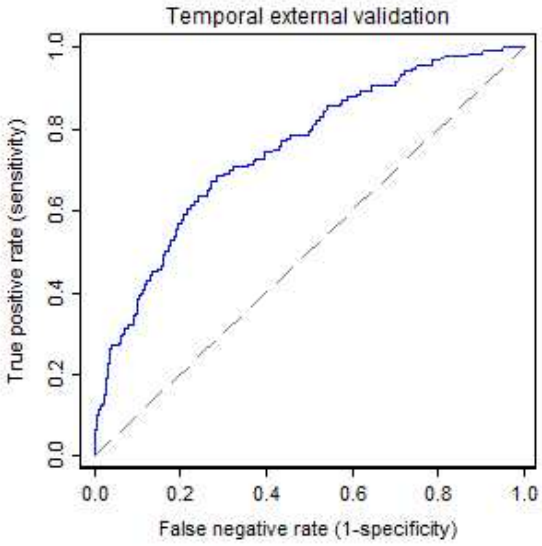
Predictors	Model derivation cohort (Sheffield) (n=1073)			Temporal validation cohort (Sheffield) (n=1087)			Broader validation cohort (Derby) (n=418)		
	aOR	95% CI	*P* value	aOR	95% CI	*P* value	aOR	95% CI	*P* value
Age, per 10 years	1.18	1.04–1.34	0.012	0.98	0.86–1.12	0.808	1.00	0.83–1.21	0.998
Prior hospitalizations, per unit	1.30	1.17–1.45	<0.001	1.17	1.05–1.31	0.005	1.05	0.91–1.20	0.511
Charlson comorbidity score, per unit	1.11	1.00–1.23	0.045	1.52	1.37–1.68	<0.001	1.66	1.44–1.90	<0.001
Mode of delivery									
Community nurse[a]	1.00	—	—	1.00	—	—	NA	—	—
Self/carer administration	1.28	0.62–2.64	0.500	0.96	0.50–1.85	0.911	NA	—	—
Infusion centre	1.61	0.90–2.89	0.108	1.37	0.76–2.45	0.298	NA	—	—
Concurrent IV antimicrobial therapy	1.89	1.03–3.47	0.041	1.18	0.61–2.30	0.622	0.78	0.25–2.44	0.671
Indication for OPAT									
Endovascular infection	1.62	0.67–3.88	0.283	0.50	0.17–1.49	0.213	1.75	0.62–4.94	0.291
Respiratory disease	0.71	0.26–1.96	0.513	1.66	0.72–3.82	0.237	2.51	1.02–6.19	0.046
Urogenital infection	1.21	0.54–2.69	0.643	1.43	0.67–3.06	0.359	NA	—	—
Bone and joint infection	0.96	0.49–1.90	0.916	0.67	0.36–1.24	0.202	1.66	0.77–3.61	0.196
Skin and soft tissue infection	0.46	0.25–0.86	0.015	0.39	0.20–0.76	0.006	3.39	0.29–39.46	0.330
Other indication	1.00	—	—	1.00	—	—	1.00	—	—
Performance on bootstrap internal validation (500 replications)									
Discrimination, *c*-statistic	0.70	0.65–0.74	—	0.78	0.74–0.82	—	0.74	0.69–0.80	—
HL goodness-of-fit statistic (df)	4.20 (8)	—	0.838	3.17 (8)	—	0.923	10.11 (8)	—	0.257
Average predicted (observed) risk, %	12.0 (11.5)	—	—	—	13.3 (12.9)	—	26.2 (25.4)	—	—
Calibration slope	0.87	0.68–1.06	—	0.93	0.78–1.07	—	0.91	0.68–1.15	—
Calibration-in-the-large	0.00	−0.20, 0.21	—	0.00	−0.21, 0.22	—	0.00	−0.26, 0.27	—
Performance of original model on external validation									
Discrimination, *c*-statistic	—	—	—	0.75	0.71–0.79	—	0.67	0.61–0.73	—
HL goodness-of-fit statistic (df)	—	—	—	10.79 (8)	—	0.214	26.61 (8)	—	<0.001
Average predicted (observed) risk, %	—	—	—	11.4 (12.9)	—	—	17.1 (25.4)	—	—

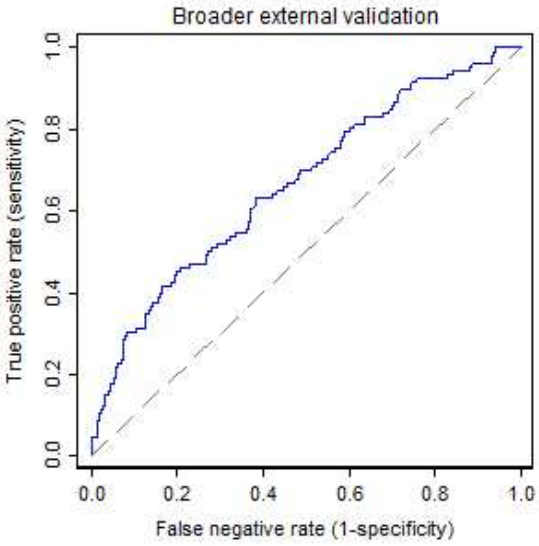
Calibration slope	—	—	—	1.05	0.84–1.26	—	1.01	0.65–1.37	—
Calibration-in-the-large	—	—	—	0.16	−0.03 to 0.35	—	0.54	0.31–0.77	—
Performance of recalibrated original model on external validation									
Discrimination, *c*-statistic	—	—	—	0.75	0.71–0.79	—	0.67	0.61–0.73	—
HL goodness-of-fit statistic (df)	—	—	—	7.07 (8)	—	0.529	3.80 (8)	—	0.875
Average predicted (observed) risk, %	—	—	—	12.9 (12.9)	—	—	26.2 (26.2)	—	—
Calibration slope	—	—	—	1.05	0.84–1.26	—	1.01	0.65–1.37	—
Calibration-in-the-large	—	—	—	0.00	−0.19–0.19	—	0.00	−0.23–0.23	—

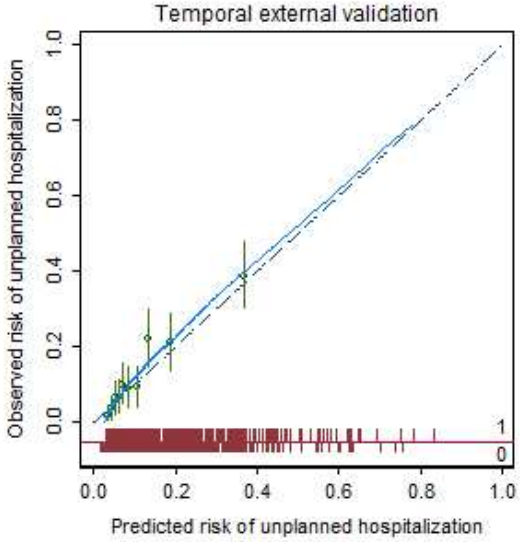
aOR, adjusted OR; df, degrees of freedom; HL, Hosmer–Lemeshow. NA indicates that in the broader validation (Derby) cohort, mode of delivery and urogenital infection could not be included as predictors because of complete separation (very few or no subjects with urogenital infection, self/carer administration or treated via infusion centre, none of whom required hospitalization).

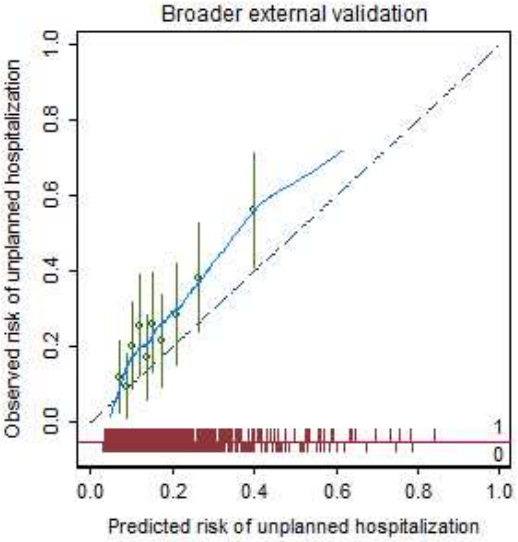
[a]Community nurse refers to administration of antimicrobial therapy in a patient's home by a nurse.

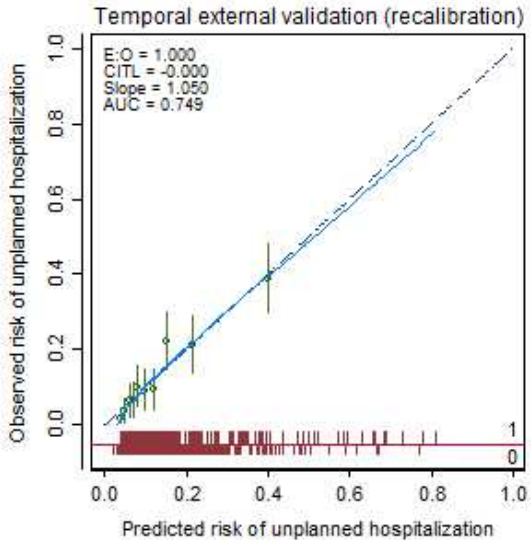
Figure 1. ROC curves and calibration plots of the multivariable prediction model for 30 day unplanned hospitalization when applied to external validation cohorts. On the ROC curves, the diagonal line indicates complete absence of discriminative ability. On the calibration plots, the smoothed line shows the agreement between predicted and observed probabilities of 30 day unplanned hospitalization. The dashed diagonal line indicates perfect calibration. The circled points represent observed proportions of unplanned hospitalization in decile groups of predicted risks, with vertical lines representing 95% CIs. The spike plot on the *x*-axis summarizes the density of patients in the range of predicted risks of unplanned hospitalization.


Temporal external validation
True positive rate (sensitivity)
False negative rate (1-specificity)
0.0
0.2
0.4
0.6
0.8
1.0


Broader external validation
True positive rate (sensitivity)
False negative rate (1-specificity)
0.0
0.2
0.4
0.6
0.8
1.0


Temporal external validation
Observed risk of unplanned hospitalization
Predicted risk of unplanned hospitalization
1
0
0.0
0.2
0.4
0.6
0.8
1.0


Broader external validation
Observed risk of unplanned hospitalization
Predicted risk of unplanned hospitalization
1
0
0.0
0.2
0.4
0.6
0.8
1.0


Temporal external validation (recalibration)
E:O = 1.000
CITL = -0.000
Slope = 1.050
AUC = 0.749
Observed risk of unplanned hospitalization
Predicted risk of unplanned hospitalization
1
0
0.0
0.2
0.4
0.6
0.8
1.0

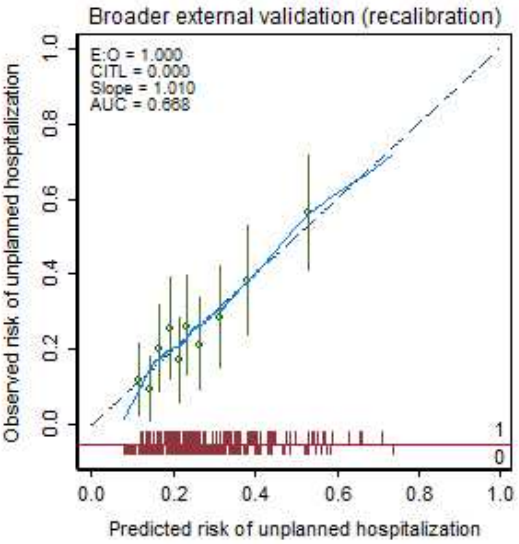

Broader external validation (recalibration)
E:O = 1.000
CITL = 0.000
Slope = 1.010
AUC = 0.668
Observed risk of unplanned hospitalization
Predicted risk of unplanned hospitalization
1
0
0.0
0.2
0.4
0.6
0.8
1.0

Figure 2. NB curves of the multivariable prediction model for 30 day unplanned hospitalization in patients receiving OPAT when applied to external validation cohorts. Solid blue line represents the NB when using the risk prediction model; dashed line represents the NB when all patients are administered therapy as inpatients ('intervention for all'); dotted horizontal line represents the NB with current practice for selecting patients for OPAT ('intervention for none').