Supplementary Material

Supplementary Methods

S1. Rationale for covariate selection and model complexity

Because only 26 cases of *Clostridioides difficile* infection (CDI) were observed among 113 procedures (25 among the 112 procedures with a known transplant date), the number of covariates that could be reliably estimated was limited. Following the commonly used rule of thumb of approximately 10 events per variable (EPV) for logistic regression, the available events supported at most 2-3 covariates (25 events / 10 ≈ 2.5). The multivariable model was therefore restricted to three clinically pre-specified covariates: diagnosis (lymphoma vs. multiple myeloma), fluoroquinolone prophylaxis (FQP) status, and calendar year of transplantation, corresponding to an events-per-variable ratio of approximately 8.3. Age, sex, transplant number, renal function and duration of grade IV neutropenia, presented descriptively in Table 1, were not entered into the multivariable model because of this sample-size constraint, not because they were assumed to be clinically irrelevant.

S2. Firth's penalized logistic regression

An events-per-variable ratio below 10, combined with a small overall sample, increases the risk of separation and of biased, unstable maximum-likelihood estimates. Firth's penalized likelihood approach was therefore used instead of standard maximum-likelihood logistic regression. This method adds a Jeffreys-invariant penalty term to the log-likelihood, which reduces small-sample bias in the regression coefficients and yields finite estimates even in the presence of quasi-complete separation. Wald-type 95% confidence intervals were derived from the penalized model. The model was fitted with diagnosis, FQP status and calendar year of transplantation (centered on the sample mean and entered as a continuous variable, in years) as covariates, coded as: diagnosis, lymphoma = 1, multiple myeloma = 0; FQP, yes = 1, no = 0.

The model was implemented in Python (NumPy/SciPy) using Newton-Raphson iteration with the standard Firth bias-correction term based on the diagonal of the hat matrix (Heinze & Schemper, Stat Med 2002). Given the exploratory nature of the analysis and the small number of events, results should be interpreted as hypothesis-generating; independent verification of these estimates using a validated statistical package (e.g., the logistf package in R, or the firthlogist package in Python) is recommended before the estimates are relied upon outside the context of this study.

Crude odds ratios for diagnosis and FQP in Table 3 are based on the full cohort (n = 113, 26 events); the crude odds ratio for calendar year and all adjusted (Firth) estimates are based on the 112 procedures with a recorded transplant date (25 events), because calendar year could not be computed for one procedure with a missing transplant date. Diagnosis and FQP status were both available for that procedure, which is why the crude estimates for these two variables use the full n = 113.

S3. Sensitivity analysis restricted to toxin-positive cases

CDI was defined in the primary analysis as a positive enzyme immunoassay result (glutamate dehydrogenase [GDH] antigen and/or toxin A/B) in a patient with diarrhea and a compatible clinical picture, in the absence of nucleic acid amplification testing. Because GDH-positive, toxin-negative results may represent colonization rather than true infection, a sensitivity analysis was performed in which CDI was redefined as a positive toxin A/B result only (toxin-positive, irrespective of GDH result). Of the 26 CDI cases in the primary analysis, 20 were toxin-positive; the remaining 6 (GDH-positive/toxin-negative in 5, inconclusive in 1) were excluded from both the numerator and the denominator of the affected subgroup in this sensitivity analysis, so that the comparison is between toxin-positive cases and procedures without any evidence of CDI. Denominators were adjusted accordingly for each subgroup (Table S1).

S4. Same-era (2023) sensitivity analysis

Because FQP status was strongly associated with calendar year (point-biserial correlation r = −0.59 with year of transplantation), a sensitivity analysis was performed restricted to procedures performed in 2023, the only calendar year in which both FQP (n = 11) and non-FQP (n = 27) procedures were represented in comparable numbers. Within this subgroup, the crude odds ratio for FQP was recalculated using the same CDI case definition as the primary analysis. This analysis was not adjusted for additional covariates because of the small subgroup size and is presented as a descriptive check on the direction of the FQP association rather than as an independent causal estimate.

S5. Confidence intervals for proportions

Wilson score 95% confidence intervals were used for all reported CDI incidence proportions (Table 2), because this method retains better coverage than the normal approximation when proportions or sample sizes are small.

S6. Software and reproducibility

All analyses were performed in Python 3 using the pandas, NumPy and SciPy libraries. Analysis code and derived summary tables are available from the corresponding author upon reasonable request.

Supplementary Tables

Table S1. Sensitivity analysis restricted to toxin-positive cases, by subgroup

Subgroup	Toxin-positive / adjusted N	Excluded (non-toxin-positive CDI)	% toxin-positive
Lymphoma (BeEAM)	7/17	1	41.2
Multiple myeloma (melphalan)	13/90	5	14.4
FQP	12/80	3	15.0
Non-FQP	8/27	3	29.6
Total	20/107	6	18.7

Adjusted N = group total minus the number of non-toxin-positive CDI cases in that group (GDH-positive/toxin-negative or inconclusive), which were excluded from both numerator and denominator. Fisher's exact test: lymphoma vs. myeloma, p = 0.017; FQP vs. non-FQP, p = 0.150

Table S2. Same-era (2023) sensitivity analysis of CDI incidence by FQP status

Group (2023 only)	CDI / n	CDI incidence, %	Crude OR (95% CI)
FQP	2/11	18.2	0.44 (0.08–2.50)
Non-FQP	9/27	33.3	reference

Fisher's exact test, p = 0.45. Analysis restricted to procedures performed in 2023, the only calendar year with a substantial number of procedures in both FQP and non-FQP groups; not adjusted for additional covariates due to the limited subgroup size

Table S3. Diagnostic test results among the 26 CDI cases

Test result	n	%
GDH-positive and toxin A/B-positive	19	73.1
Toxin A/B-positive, GDH-negative	1	3.8
GDH-positive, toxin A/B-negative	5	19.2
Inconclusive	1	3.8
Total	26	100.0

Toxin A/B-positive (rows 1–2, n = 20) defines the toxin-positive subgroup used in the sensitivity analysis (Table S1)

Table S4. Firth logistic regression model – full output

Term	Coefficient (log OR)	SE	OR (95% CI)	p
Intercept	−0.49	0.50	0.61 (0.23–1.62)	0.324
Lymphoma (vs. myeloma)	1.45	0.58	4.25 (1.37–13.22)	0.012
FQP (vs. none)	−1.43	0.66	0.24 (0.06–0.87)	0.030
Year of transplantation (centered, per year)	−0.43	0.35	0.65 (0.32–1.30)	0.223

n = 112 procedures, 25 CDI events. SE, standard error (Wald); OR, odds ratio; CI, confidence interval