Health policy 161 (2025) 105416

Contents lists available at ScienceDirect

Health policy

journal homepage: www.elsevier.com/locate/healthpol

Financial penalty associated with a decline in hospital-acquired complications in Australia☆

Luke Slawomirski [a,*], Petr Otahal [a], Martin Hensher [a], Julie Campbell [a], Stephanie Newell [b], Barbara de Graaff [a]

[a] Menzies Institute for Medical Research, University of Tasmania,17 Liverpool St (Private Bag 23), Hobart 7000 Tasmania, Australia
[b] Australian Patient Safety Network, Adelaide, South Australia, Australia

ARTICLE INFO

Keywords:
Pay-for-performance (p4p)
Financial penalties
Financial incentives
Patient safety
Adverse events
Iatrogenic harm
Inpatient

ABSTRACT

Background: Adverse events during hospital care are a global concern. The evidence for addressing unsafe acute care using pay-for-performance (p4p) is inconclusive.
Objective: To examine association between the introduction of a financial penalty on 1 July 2018 and the prevalence of 13 high-priority hospital-acquired complications (HACs) in Australian public hospitals.
Methods: Administrative data on every Australian public hospital separation (age >17 years) between 1 January 2014 and 30 June 2021 was used to analyse changes in quarterly HAC prevalence (per 1000 multi-day separations), standardized to the study population, using two interrupted time series methods: generalized least squares (GLS) with autoregressive moving average (ARMA) errors, and a Bayesian structured time series.
Results: Just under 20 million separations took place over the study period with 947,057 (4.7%) (mean age 69 (SD: 18), 48% female) recording at least one HAC and 1,263,646 HACs overall. Our GLS model estimated a decline of 17% (95% CI 12 – 22%) in HAC prevalence associated with the introduction of the penalty. The Bayesian model estimated a 26% (23 – 29%) decline. Most of the decline occurred during a 12-month roll-in period. Results suggest that 98,970 fewer inpatients experienced a HAC from 1 July 2018 to 30 June 2021 compared to the modelled counterfactual.
Conclusions: Implementation of a financial penalty was associated with a substantial decline in HACs. Few other p4p policies have been associated with reductions in inpatient harm. Future research should examine local HAC trends and investigate what other factors may have contributed to the change.

Research in context

What is already known about the topic?

International attempts to use financial incentives to improve safety of hospital care have produced mixed results. An Australian policy that deducts the additional cost of hospital-acquired complications (HACs) from public hospital funding was introduced in 2018 but has not yet been formally evaluated.

What does this study add to the literature?

This study found a meaningful reduction in HAC rates associated with the introduction of the policy in Australian public hospitals, although the decline is most likely the result of several factors, policies and improvement activities taking place around the same time.

What are the policy implications?

Financial incentives can complement other efforts to achieve policy objectives. Attention should be paid to how schemes are designed and implemented, especially from a stakeholder perspective.

☆ This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.
* Corresponding author.
E-mail addresses: Lukasz.Slawomirski@utas.edu.au (L. Slawomirski), Petr.Otahal@utas.edu.au (P. Otahal), Martin.Hensher@utas.edu.au (M. Hensher), Julie.Campbell@utas.edu.au (J. Campbell), newellsteph@gmail.com (S. Newell), Barbara.degraaff@utas.edu.au (B. de Graaff).

https://doi.org/10.1016/j.healthpol.2025.105416
Received 2 April 2025; Received in revised form 6 August 2025; Accepted 9 August 2025
Available online 11 August 2025

1. Background

Adverse events continue to exert a substantial and unnecessary burden on patients and healthcare systems [43]. Safety lapses during inpatient care are especially burdensome [18,19,31,38]. Policies to reduce inpatient harm include pay for performance (p4p), where financial incentives are used to align clinical and corporate risk with the goal of stimulating greater organizational action on patient safety [37]. P4p draws on principles of neoclassical economics, in particular rational choice theory [11]. A common approach is to subtract from reimbursement the additional costs generated by an adverse event. This *de facto* financial penalty is grounded in behavioral economics, specifically prospect theory, which suggests that individual agents tend to value financial losses more than gains [34]. Less is known, however, about the impact on teams and organizations, and evidence of sustained improvement in outcomes associated with financial penalties remains sparse. Questions remain regarding the optimal size of the incentives, whether financial incentives alone are enough to stimulate organizational change, and whether sufficient consideration has been given to the way these schemes have been implemented particularly from a stakeholder management perspective [37].

A national financial penalty policy targeting a subset of adverse events termed 'hospital-acquired complications' (HACs) was introduced across Australia's 692 public hospitals on 1 July 2018 [4,25]. The Australian Commission on Safety and Quality in Health Care (ACSQHC)—the national agency responsible for developing and specifying the HAC list—defines a HAC as "a patient complication for which clinical risk mitigation strategies may reduce (but not necessarily eliminate) the risk of that complication occurring". HACs are a selection of diagnoses that feature in hospital administrative data and are coded '*not* present on admission'. They were selected based on their relatively frequency, and their discernibly negative impact on patients and health services [2]. Thirteen of the 16 HAC categories currently specified are subject to the financial penalty (Box 1).

Two studies have examined HAC rates in Australian public hospitals using administrative data [17] studied three years of data from 38 public hospitals in South Australia and Victoria, observing that 9.7 % of multi-day episodes recorded at least one HAC. Edwards et al. [18] observed a rate of 9.3 % in four South Australian public hospitals over five years. Previously, Ehsani et al. [19] observed that 19 % of multi-day episodes in Victorian public hospitals recorded at least one hospital-acquired diagnosis—although these included all conditions, not just the ones that would eventually be defined as HACs.

A recent systematic review of p4p initiatives targeting inpatient safety found inconclusive evidence for their effectiveness [37] but did not include the Australian policy, which has not been evaluated in the academic literature. In this study we used national admitted patient data to assess changes in HAC rates associated with the introduction of the financial penalty across Australian public hospitals.

2. Methods

2.1. Study design, population and data sources

This observational study used administrative data collected from all Australian public hospitals to examine the impact of the penalty. We used an interrupted time series design. Our study population (for which we had record-level data) was every inpatient aged 18 years and over separated from an Australian public hospital with at least one diagnosis coded '*not* present on admission' between 1 January 2014 and 30 June 2021 ($n = 4536{,}575$). We included only multi-day separations (length of stay > 1 bed day) because preliminary analysis suggested that these account for 99 % of HACs (multi-day separations make up over 80 % of bed days in Australian public hospitals [7]). These data were provided by the Independent Hospital and Aged Care Pricing Authority (IHACPA) and originate from the Admitted Patient Care National Minimum Dataset (APC—NMDS)—administrative data abstracted from medical records by clinical coding professionals [6]. The data included variables on demographics, separation dates, up to 99 secondary diagnoses (using ICD-10-AM 8th-11th editions) with their corresponding Condition Onset Flag (COF) (COF1: *not* present on admission (i.e. hospital-acquired); COF2: present on admission (POA); COF 9: not specified); and up to fifty procedure codes using the Australian Classification of Health Interventions (ACHI). We used the ACSQHC algorithm version 3.1 to identify HACs in our study population [2]. This algorithm uses ICD10-AM, COF and ACHI codes to identify and extract HACs from the APC—NMDS. Our denominators – total quarterly public hospital separations over the study period ($N = 44{,}014{,}518$) as well as mean/median length of stay (LOS), mean/median age and percentage female patients for each study quarter – we obtained separately from the Australian Institute of Health and Welfare (AIHW), which maintains the APC—NMDS [6].

Box 1
HAC categories, their percentage of all HACs counted over the entire study period, and inclusion status in financial penalty policy.

HAC1	**Pressure injury** (1.9 %)	**Included**
HAC2	**Falls resulting in fracture or intracranial injury** (1.4 %)	**Included**
HAC3	**Hospital-acquired infection** (30 %)	**Included**
HAC4	**Surgical complications requiring unplanned return to theatre** (13 %)	**Included**
HAC5	*Unplanned intensive care admission*[1]	
HAC6	**Respiratory complications** (8 %)	**Included**
HAC7	**Venous thromboembolism** (2.6 %)	**Included**
HAC8	**Renal failure** (0.5 %)	**Included**
HAC9	**Gastrointestinal bleeding** (3.2 %)	**Included**
HAC10	**Medication complications** (3.4 %)	**Included**
HAC11	**Delirium** (13 %)	**Included**
HAC12	**Incontinence** (1.9 %)	**Included**
HAC13	**Endocrine complications** (7.1 %)	**Included**
HAC14	**Cardiac complications** (13 %)	**Included**
HAC15	*Third- and fourth-degree perineal laceration during delivery*	
HAC16	*Neonatal birth trauma*	

2.2. The financial penalty policy

The financial penalty policy, although introduced in July 2018, initially arose from the Australian 2011 National Health Reform Agreement [15]. It was jointly developed by the IHACPA and ACSQHC. The six-year process to define and specify HACs, then design and introduce the policy, included consultation with consumer representatives, clinical leaders and policy makers from all Australian States and Territories. It comprised a proof-of-concept, a pilot, and a 'shadow year' in 2017–18 during which the penalty was calculated and the expected funding adjustments shared with States and Territories for review, verification and feedback for the final design of the model [2,25]. The penalty is calculated at the record level and is adjusted using a model that considers individual patient characteristics including age and comorbidities (Supplementary file: S1).The resulting downward adjustment to total reimbursement (i.e. the penalty) has been relatively modest so far, totaling approximately one percent of overall hospital funding during the policy's first three years [24].

2.3. Outcomes

Our primary outcomes were the change in quarterly HAC prevalence (number of HACs per 1000 separations) and in the quarterly HAC episode rate (number of separations featuring at least one HAC per 100 separations). As the payment adjustment was rolled out simultaneously across all Australian public hospitals, no regions or hospitals could serve as comparisons. Instead, we used prevalence of hospital-acquired diagnoses not classified as a HAC—henceforth *other hospital-acquired diagnoses*—for this purpose. Our secondary outcomes were the difference between expected and observed HACs and HAC episodes during the post-intervention period calculated using ITS model estimates .

We also endeavored to detect changes in coding practice that might have been deployed to reduce the number of HACs recorded. Specifically, we examined HAC diagnoses coded as 'present on admission' (COF2) and 'not specified' (COF9)—which can be used to mask true HACs—and ICD-10-AM codes that can be used as an alternative to describe iatrogenic conditions and diagnoses: T80–88.9 'complications of surgical and medical care not elsewhere classified'; Y40-Y84.9 'complications of medical and surgical care'); and Y95 ('nosocomial condition')—henceforth *procedural complication codes*. We excluded these from the count of diagnoses coded 'not present on admission'.

2.4. Statistical analysis

We performed descriptive analyses to summarize continuous data with mean and standard deviation or median and interquartile range, where appropriate. We segmented the study period into three phases: pre-introduction phase (Jan 2014-Jun 2017: study quarters 1–14); shadow year (Jul 2017-Jun 2018: quarters 15–18); and the policy phase (Jul 2018-Jun 2021: quarters 19–30). We standardized all outcomes to our December 2017 quarter study population using the Charlson Comorbidity Index (CCI) scores with age-adjustment calculated at the record-level [14]. We used the ICD-10 coding algorithm from Quan et al. [32] and patient age to create six CCI strata based on the distribution of CCI scores in the study population ('0, '1–3′, '4′, '5′, '6–8′ and '9 and above').

We chose interrupted time series (ITS) analysis as our method. ITS assesses changes in the level and/or gradient of a dependent time series and is the preferred quasi-experimental design to assess the impact of large-scale policy interventions [10,40]. ITS has several advantages over a comparison of pre- and post-intervention grand means. It provides more robust insights into baseline pre-implementation trends and counterfactual rates, it can separate post-implementation period into short-term versus long-term effects that may have different aetiologies, provide an estimate of causal effect, and it enables more robust adjustment for autocorrelation and heteroskedasticity.

We used two complementary ITS approaches: a generalized least squares (GLS) regression model with autoregressive moving average (ARMA) errors, and a Bayesian structural time series approach[26,35,42]. A more detailed description of both is provided in Supplementary file: Statistics. We used the estimates from the Bayesian models to calculate the difference between observed and expected HACs and HAC episodes.

We used July 2018 (study quarter 19) as the principal interruption point in our models. However, we also modelled the time series without the shadow year (study quarters 15–18) thus treating that as a "roll-in" period. We did not adjust models because (1) our outcomes were already standardized, and (2) the covariates at our disposal (quarterly mean LOS and proportion female for all inpatient separations) varied little across the study period. We conducted ITS analyses on HACs1–14 combined and provide descriptive analyses for individual HACs. We considered a two-sided alpha of <0.05 to be significant.

We performed two sensitivity analyses. The first tried to account for Covid-19, which arrived in Australia in early 2020 (study quarters 25–26) and is likely to affect our estimates because: (1) all separations with a Covid-19 principal or secondary diagnosis were exempted from the penalty [22], (2) national and local responses to hospital admission policy (12 % fewer adult hospital admissions occurred in April-June 2020 compared to the four-quarter average immediately before and after), and (3) the impact of Covid-19 on frontline care [8,27,28]. We therefore replaced the observed outcomes in quarters 26–30 with those of (pre-Covid) quarter 25. Second, we tried to further account for temporal changes in the aggregate public hospital casemix by using national public hospital funding as the denominator. Australian public hospital services are predominantly funded on casemix—aggregate funding therefore reflects complexity as well as volume of care. Moreover, funding is reported publicly in monthly increments [30], enabling us to generate 90 data points over our study period. We adjusted for inflation by indexing each funding figure to the 2014–15 National Efficient Price [23]. One obvious limitation of this approach is that, after July 2018, funding was subject to the downward HAC adjustment, which would deflate our denominator (thus inflating HAC prevalence). However, as the final adjustments ended up being small (about 1 % of aggregate funding) we did not re-inflate the funding figures post-July 2018. Also, the monthly funding figures are erratic, owing to periodic reconciliation. We addressed this issue by using the raw amounts, as well as a six-month simple moving average (SMA) to calculate HAC rates and tested both in our models.

All analyses were conducted using R version 4.3.3 [33]. Figures and tables were created in Microsoft Excel and R. The study was approved by the University of Tasmania Research Ethics Unit with a waiver of consent because all data were de-identified and analyzed in a secure virtual environment. We followed the Strengthening Reporting of Observational Studies in Epidemiology (STROBE) Statement Guidelines including the checklist [41].

3. Results

3.1. Study population

A total of 44,014,518 adult public hospital separations occurred over the study period, of which 19,951,833 were multi-day separations (Fig. 1). The mean age of multi-day patients was 58 (SD: 23) years and 55 % were female. Of the multi-day separations, 4229,144 (21 %) had at least one diagnosis coded 'not POA'. The mean age was 57 (SD: 22) years, median age-adjusted CCI score was 3 (IQR: 0–5) and 62 % were female (Table 1). We counted 947,057 HAC episodes (at least one recorded HAC). The mean age of this cohort was 69 (SD: 18) years, 48 % were female and the median CCI score was 4 (IQR: 3–7). We observed little temporal change in the descriptive characteristics of these cohorts across the study period (Supplementary file: S2). We counted a total of 1263,646 HACs included in the penalty policy (Table 1). The most

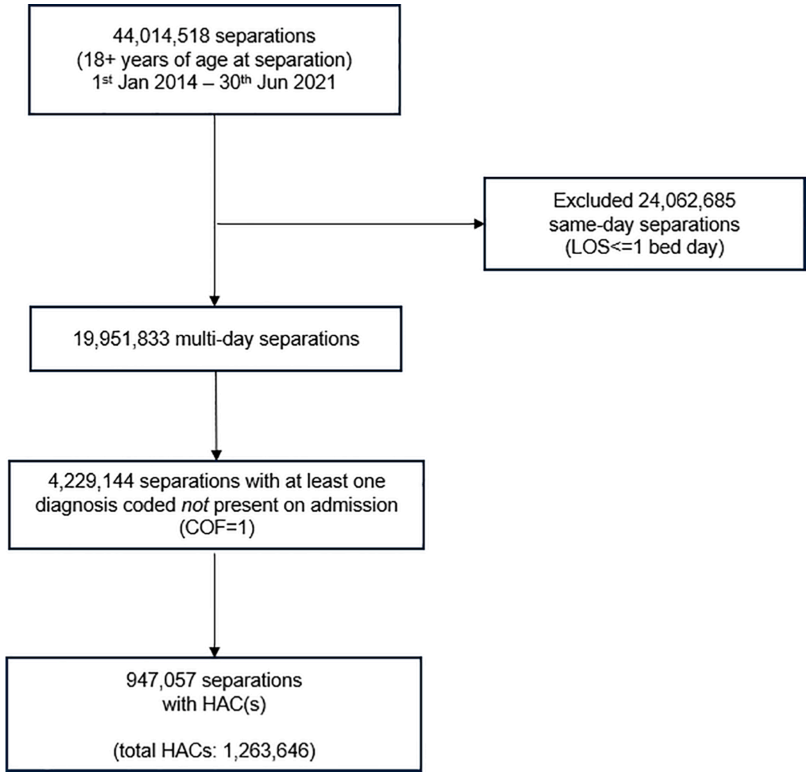

Fig. 1. Strengthening the reporting of observational studies in epidemiology (STROBE) diagram of exclusions used in the study population.

Table 1
Study population.

		All adult multi-day separations	Separations with at least one diagnosis coded 'not present on admission' (i.e. other hospital-acquired diagnoses)	Separations with one or more HACs1–14	Total HACs counted
Entire study period	N	19,951,833	4229,144	947,057	1263,646
	mean age (SD) years	58 (23)	57 (22)	69 (18)	…
	% female	55 %	62 %	48 %	…
	median CCI[a] (IQR)	..	3 (0–5)	4 (3–7)	….
Jan 2014-Jun 2017 (pre-introduction phase)	N	8968,877	1927,246	485,443	657,114
	mean age (SD) years	57 (60)	57 (60)	69 (73)	…
	% female	55 %	62 %	48 %	…
	median CCI[a] (IQR)	…	3 (0–5)	4 (3–7)	…
Jul 2017-Jun 2018 (shadow year)	N	2735,431	603,047	133,616	178,491
	mean age (SD) years	58 (60)	57 (61)	69 (73)	…
	% female	55 %	62 %	48 %	…
	median CCI[a] (IQR)	…	3 (0–5)	4 (3–7)	…
Jul 2018-Jun2021 (policy phase)	N	8247,525	1698,851	327,998	428,041
	mean age (SD) years	58 (61)	57 (60)	69 (73)	…
	% female	55 %	62 %	48 %	…
	median CCI[a] (IQR)	…	3 (0–5)	4 (3–7)	…

a. age-adjusted Charlson Comorbidity Index score.

frequent HACs were HAC3: infections (30 %), HAC14: cardiac complications (13 %), HAC4: surgical complications (13 %) and HAC11: delirium (13 %) (Supplementary file: S3).

3.2. Outcomes

Over the entire study period HAC prevalence was 64 per 1000 separations and the HAC episode rate was 4.7 %. The time series for both primary outcomes are presented in Fig. 2. Using the main interruption point (July 2018), our GLS models estimated a decline of 17 % (95 % CI 12 – 22 %) in quarterly HAC prevalence and a decline of 15 % (11 – 19 %) in HAC episode rates. The pre- and post-interruption slope was downward for both outcomes but no change in gradient was observed. Both models performed adequately. The Bayesian models estimated declines of 26 % (23 – 29 %) in HAC prevalence, and 24 % (21 – 27 %) in HAC episode rates, although the prevalence model performance may

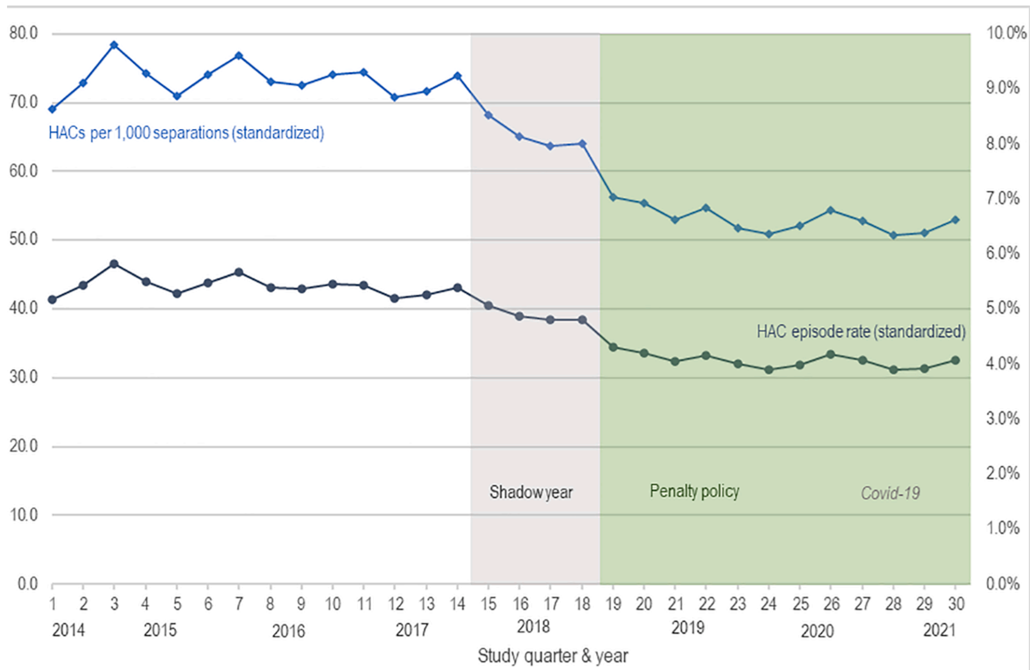

Fig. 2. Standardized HAC prevalence and HAC episode rate, Australian public hospitals (Jan 2014 – Jun 2021).
Quarterly standardized HAC incidence (blue line, left vertical axis) and standardized HAC episode rates (black line, right vertical axis) over the study period.

potentially overstate the effect magnitude (Table 2). We estimate that 148,455 (123,713 – 173,198) fewer HACs, and 98,970 (82,475 – 115,465) fewer HAC episodes, were recorded from July 2018 to the end of the study period than would have been under the counterfactual scenario.

With the shadow year removed, GLS model estimated a decline of 26 % (23 – 28 %) in HAC prevalence as well as a slope change from zero to −0.40 (−0.57 – −0.23). However, the model's Box-Ljung test was borderline ($p = 0.047$). The GLS model for HAC episode rates performed adequately, estimating a decline of 22 % (20 – 24 %) but no significant change in the slope. The Bayesian models estimated declines of 29 % (31 – 27 %) and 27 % (24 – 29 %) respectively, with both models exhibiting good fit (Table 2).

3.2.1. *Comparator*

The GLS models examining changes in the quarterly prevalence of other hospital-acquired diagnoses performed inadequately, as did the Bayesian model using the July 2018 interruption. With the shadow year removed, the Bayesian model fitted well, estimating a decline of 5.1 % (0.96 – 8.9 %) (Table 2, Supplementary file S4).

3.2.2. *Changes in coding practice*

For HAC diagnoses coded POA, the GLS model using the July 2018 interruption performed inadequately. With the shadow year removed the GLS model performed well, estimating a change in slope from 1.14 (0.68 – 1.6) to −0.81 (−1.4 – −0.23) but no change in level. The Bayesian models estimated no change using both interruption points (Table 2, Supplementary file S5). For the procedural complication codes, the GLS model estimated a decline of 4.9 % (0.9 - 9.0 %) and a change in slope from −0.20 (−0.31 – −0.10) to 0.4 (0.21 – 0.59) after July 2018. Removing the shadow year, the model estimated a decline of 8.5 % (4.1 – 13 %) and a slope change from zero to 0.37 (0.16 – 0.59). The Bayesian models estimated no change after July 2018, and a decline of 4.4 % (2.0 – 6.8 %) with the shadow year removed (Table 2).

3.3. *Sensitivity analyses*

Substituting pre-Covid-19 prevalence for the remaining pandemic-affected, quarters of the study period produced similar results to the principal analysis. For HACs per $1 million of monthly hospital funding, our GLS model estimates were also similar to the principal analysis, whilethe Bayesian models estimated a decline of 42 % (38 – 45 %) using the July 18 interruption, and a decline of 41 % (38 – 44 %) with the shadow year removed. Models using the raw monthly funding and the SMA funding figures produced nearly identical results, but the latter performed poorly due to autocorrelation. We therefore report the estimates derived from the raw funding figures (Table 2, Supplementary file: S6).

4. Discussion

This large observational study drawing on all Australian public hospital separations between January 2014 and June 2021 found that the introduction of a small financial penalty was associated with a meaningful decline in the prevalence of HACs targeted by the scheme. The estimated decline ranged from 16 % to 42 % depending on models and denominators used. We observed a modest decline in hospital-acquired diagnoses that did not attract the penalty, suggesting a possible positive spillover effect. Our observed pre-policy HAC episode rate (5.4 %) was lower than that observed by two Australian studies examining HAC prevalence (9.7 % and 9.3 %)—which used data from a sample of hospitals, included HAC15, and used an older algorithm to identify HACs [17,18]—but closer to the rates observed by IHAPCA if sameday separations are included [21].

We must be cautious about attributing the decline to the penalty alone. For one, it was among several concurrent national efforts to improve patient safety in acute care (pre-intervention rates were modestly declining in the pre-introduction phase). A national focus on monitoring and reducing hospital-acquired infections have been in play for close to two decades [3], for example, and the National Safety and Quality Health Service (NSQHS) Standards – Australia's first national hospital standards – were introduced in 2013 [1].

The observed decline was in the level, not the slope, of quarterly HAC rates. This finding is open to interpretation. A micro-economic perspective may be that the level of attention given to patient safety is determined inter alia by an implicit, baseline equilibrium between the

Table 2
Study results from generalized least squares (GLS) and Bayesian ITS models[a].

		Main outcomes		Comparator	Coding practice		Sensitivity analyses	
		HACs per 1000 separations	HAC episodes per 100 separations	Other hospital-acquired diagnoses not classified as HACs, per 1000 separations	HAC diagnoses coded POA per 1000 separations	Procedural complication codes per 1000 separations	HACs per 1000 separations Covid-adjusted[b]	HACs per $1 M hospital funding (MONTHLY)
July 2018 interruption	**GLS model**							
	Pre-interruption slope (95 % CI)	−0.51 (−0.78, −0.23)	−0.040 (−0.055, −0.025)	1.26 (−0.20, 2.72)	0.54 (−1.48, 2.56)	−0.20 (−0.31, −0.10)	−0.51 (−0.79, −0.22)	−0.021 (−0.041, 0.000)
	Post-interruption slope (95 % CI)	−0.39 (−0.89, 0.10)	−0.027 (−0.055, −0.0002)	0.47 (−2.08, 3.01)	−0.91 (−3.42, 1.6)	0.40 (0.21, 0.59)	−0.47 (−0.96, 0.03)	−0.020 (−0.047, 0.006)
	Level change (95 % CI)	−13.0 (−16.8, −9.1)	−0.84 (−1.01, −0.62)	−25.6 (−47.3, −3.82)	10.4 (1.7, 19.2)	−2.02 (−3.68, −0.37)	−12.8 −16.4, −9.1	−0.75 (−1.22, −0.27)
	Relative change (95 % CI)	−17 % (−22, −12)	−15 % (−19, −11)	−11 % (−21, −2)	16 % (3.4, 31)	−4.9 % (−9.0, −0.9)	−17 % (−20, −11)	−15 % (−24, −5)
	Modelled pre-interruption mean (95 % CI)	77 (74, 80)	5.8 (5.6, 5.9)	226 (210, 242)	64 (−480, 608)	41 (40, 42)	77.1 73.9, 80.3	5.0 (−16, 26)
	Residual standard error	23.27	0.21	12.2	277	1.22	3.2	10.5
	Box-Ljung test	0.21	0.20	*0.016*	*0.002*	0.24	0.25	0.13
	Bayesian model							
	Relative change (95 % CI)	−26 % (−29, −23)	−24 % (−27, −21)	−5.2 % (−8.6, −1.7)	−1.1 % (−6.1, 4.4)	−2.9 % (−6.1, 0.5)	*−28 % (−31, −24)*	*−42 % (−45, −38)*
	Posterior tail area p-value	*0.001*	*0.001*	*0.003*	0.337	0.06	*0.001*	*0.001*
	Model fit	Limited—may overstate effect magnitude	Good	Poor—interpret with caution	Good	Good	Limited—may overstate effect magnitude	Good
		HACs per 1000 separations	HAC episodes per 100 separations	Other hospital-acquired diagnoses not classified as HACs, per 1000 separations	HAC diagnoses coded POA per 1000 separations	Procedural complication codes per 1000 separations	HACs per 1000 separations Covid-adjusted[b]	HACs per $1 M hospital funding (MONTHLY)
Shadow year quarters removed (interruption at July 2017)	**GLS model**							
	Pre-interruption slope (95 % CI)	−0.055 (−0.187, 0.077)	−0.013 (−0.022, 0.004)	2.91 (1.94, 3.88)	1.14 (0.68, 1.6)	−0.99 (−0.27, 0.07)	−0.049 (−0.19, 0.096)	−0.010 (−0.018, −0.002)
	Post-interruption slope (95 % CI)	−0.40 (−0.57, −0.23)	−0.027 (−0.039, −0.015)	0.78 (−0.57, 2.13)	−0.81 (−1.4, −0.23)	0.37 (0.16, 0.59)	−0.44 (−0.63, −0.26)	−0.021 (−0.031, −0011)
	Level change at interruption (95 % CI)	−19.2 (−20.8, −17.6)	−1.24 (−1.35, −1.13)	−38.7 (−52.4, −25.0)	−2.49 (−7.89, 2.91)	−3.42 (−5.2, −1.65)	−19.2 (−20.1, −17.4)	−1.35 (−1.64, −1.07)
	Relative change at interruption (95 % CI)	−26 % (−28, −23)	−22 % (−24, −20)	−18 % (−24, −12)	−3.8 (−12, 4.4)	−8.5 % (−13.2, −4.1)	−26 % (−28, −23)	−27 % (−33, −22)
	Modelled pre-interruption mean (95 % CI)	75 (74, 76)	5.6 (5.5, 5.7)	217 (209, 225)	66 (62, 70)	40 (39, 42)	75 (74, 76)	4.9 (4.7, 5.1)
	Residual standard error	2.77	0.20	9.58	3.34	1.20	2.62	0.42
	Box-Ljung test	*0.047*	0.07	*0.002*	0.08	0.38	0.17	0.15
	Bayesian model							
	relative change (95 % CI)	−29 % (−31, −27)	−27 % (−29, −24)	−5.1 % (−8.8, −0.96)	0.12 % (−5.2, 5.9)	−4.4 % (−6.8 %, −2.0 %)	−29 % (−31, −27 %)	−41 % (−44, −38)
	Posterior tail area p-value	*0.001*	*0.001*	*0.015*	0.47	*0.001*	*0.001*	*0.001*
	Model fit	Good	Good	Good	Acceptable	Good	Good	Good

[a] All results standardized to December 2017 quarter.
[b] HAC prevalence of quarter 25 applied to quarters 26–30.

marginal cost of HACs and the marginal cost of preventing them ([16] observed that HACs exerted a net cost on Australian hospital finances). The impending financial penalties may have prompted changes in practice to a degree that re-established this equilibrium with the expected financial losses factored in (with the changes taking place predominatly during the shadow year). It can be postulated that, ceteris paribus, HAC rates 'find' a level permitted by existing technologies, practices and resources. Hospitals treat sick patients. Even with the best mitigation strategies, things do not always go to plan, although the persistent (if modest) fall in pre- as well as post-interruption HAC rates could be interpreted as mitigation strategies having to yet reach their potential. The decline can also be a data artifact, created by changes in clinical documentation and coding with the post-intervention prevalence reflecting the lowest level permitted by the new practices. The standard of clinical coding in Australia is high [36] and we failed to observe evidence for miscoding. But there are other ways to minimize the number of HACs recorded in administrative data.

Nevertheless, our findings contrast with most of the existing literature evaluating p4p and inpatient harm [37]. The Australian policy most closely resembles the United States' Centers for Medicare and Medicaid Services (CMS) Hospital-Acquired Condition Present on Admission (HAC-PoA) program, which generated inconclusive results [12]. The mechanisms explaining our results can be debated. One criticism of HAC-PoA was that the small size of the financial penalty (<0.1 % of total hospital funding [13,29]) is insufficient to inspire organizational change, although the impact of the Australian policy on aggregate hospital funding was also modest: 1.08 % for 2017–18 (the shadow year), then 0.96 % (2018–19), 0.97 % (2019–20) and 0.90 % (2020–21) [24].

What, then, could explain the effect? Changing organizational behavior is a socio-technical endeavor that requires thoughtful change management [9]. A key reason for the gradual and inclusive development of the HACs framework (a policy 'intervention' in its own right) and the subsequent penalty was to stimulate local activity to improve safety. The development process began in 2012 (outlined previously) with all stages involving clinicians, consumers and all jurisdictions [2]. By the time the policy was 'shadowed' in 2017–18 it had not just been augured, but discussed, debated and co-designed with the Australian public hospital community. Tedesco [39] found that improvement preceded the introduction of p4p initiatives targeting hospital safety in the United States, suggesting "intense public discourses" about patient safety and the ensuing policies as a potential reason [39]. In addition, the 'bark' of the Australian policy may have been louder than its 'bite' in terms of the eventual modest size of the penalties.

There may be a temptation to increase the size of the penalty to stimulate further reduction. We would urge caution due to the risk of counterproductive behavior and unintended consequences. For example, clinicians may become overly cautious about mobilizing patients to reduce falls, thereby potentially extending length of stay and increasing the risk of other complications; penalizing hospitals that are already under-resourced could entrench and exacerbate existing problems; and the temptation to game would rise. Instead, more research on the behavioral and organizational drivers of the observed changes associated with this policy at the hospital level is advised. This may also shed light on whether efforts to reduce HACs diverted resources from other activities.

Overall, our findings suggest that p4p may complement a broader suite of policies and, as we have previously argued, that the key driver is not the economic self-interest of providers but simply greater attention drawn to patient safety outcomes [37]. Whatever the mechanism(s), we estimate that almost 99,000 fewer patients experienced a HAC than would otherwise have under a counterfactual scenario—an average of 33,000 a year. Given that a HAC typically adds about five bed days to a hospital episode [16,20], this reduction would have freed up 165,000 bed days annually, equating to just under 55,000 typical hospital admissions based on an average LOS of 3.0 bed days as reported by AIHW for 2018–19 to 2020–21 [5].

4.1. Strengths and limitations

A key strength of this study is its size. Our study captured *all* public hospital separations, removing the risk of sampling error, and covered a period of seven-and-a-half years . We standardized our data for age and comorbidities, removing temporal changes in patient structure. Our estimate of the policy effect is conservative because we excluded same-day separations (in the denominator) when calculating HAC rates. Same-day separation only contain 1 % of HACs but grew as a proportion of all separations by 4 % over the study period. This growth would have resulted in the appearance of a greater reduction in HAC rates in the post policy period. We used two distinct ITS approaches, thereby providing two stastical perspectives on the observed outcomes. We conducted thorough and well reasoned senstivity analyses to examine the robustness of the main findings.

Our study has several limitations. Firstly, we had to use quarterly intervals because we were provided with separation numbers (denominators) in quarterly increments. However, our sensitivity analysis using *monthly* aggregate hospital funding as a denominator produced similar findings. Secondly, we standardized results using the study population, for which we had the requisite record-level information, and not the entire patient population. While our cohort was large (over 4 million records, representing over 20 % of all separations) we can't be sure that temporal changes in its casemix reflect those for all inpatients. However, the consistent demographic and LOS information in our denominator numbers over the study period reassures us that overall casemix was, like our study population, relatively steady over the study period (see Supplementary file: S2). Again, our sensitivity analysis using aggregate hospital funding, which reflects care complexity as well as volume, as a denominator produced similar findings.

Using HACs as the outcome is potentially problematic for two reasons: (1) HACs are a construct whose presence (or absence) simply reflects whether certain codes appear in the APC—NMDS. While associations between HACs and endpoints such as mortality, readmission and higher costs have been established [16,18–20,31,38], strictly speaking, our study has not provided evidence that their decline is associated with any changes in these 'harder' endpoints (we could not link our data either longitudinally or with other datasets); and (2) as mentioned, they are susceptible to changes in documentation and coding. Our census approach using data from all public hospitals curbs the potential effect of isolated miscoding. Widespread miscoding may have been suggested by increased rates of HAC diagnoses coded as 'present on admission' and of procedural complications codes. We observed neither. Further, the number of codes that denote a HAC increases with each edition of ICD-10-AM (our study period covers four editions) and it is unlikely that this would reduce the number of HACs coded over time.

Another limitation is the arrival of Covid-19 in early 2020, which changed hospital practices for the final quarters of our study period. Standardizing our results to the 2017 December quarter of our study population went some way to account for this and the sensitivity analysis addressing Covid-19 did not change our findings. As discussed, lack of a comparator hospitals further limits our ability to attribute the observed declines to any one variable.

Finally, we were only authorized to report national aggregates. Results for individual jurisdictions, hospitals and patient classes are likely to differ. Such analysis would provide valuable insights into what drives improvement in patient safety and vice versa. Similarly, we focused on combined counts of HAC1–14 (time series for each HAC are provided in the Supplementary file: S3). More detailed analysis of individual HAC classes could also prove useful, potentially revealing novel ways to reduce patient harm.

5. Conclusions

In this large, observational study drawing on seven-and-a-half years of Australian public hospital separations, the implementation of a financial penalty was associated with a decline in the national rate of hospital-acquired complications. The careful and gradual fashion in which the policy was developed and implemented—rather than the penalty itself—is a plausible driver. The influence of other, separate, policies and practices cannot be ruled out. While the decline in complications is good news for Australian inpatients, over 110,000 hospital separations still recorded a HAC during the last year of our study. That amounts to about 300 patients a day. Efforts to improve the safety of inpatient care must continue. Future research should focus on identifying the drivers of the observed change in HAC rates.

CRediT authorship contribution statement

Luke Slawomirski: Writing – original draft, Visualization, Methodology, Formal analysis, Data curation, Conceptualization. **Petr Otahal:** Writing – review & editing, Validation, Methodology, Data curation. **Martin Hensher:** Writing – review & editing, Supervision, Conceptualization. **Julie Campbell:** Writing – review & editing, Supervision, Conceptualization. **Stephanie Newell:** Writing – review & editing, Visualization. **Barbara de Graaff:** Writing – review & editing, Supervision, Project administration, Conceptualization.

Declaration of competing interest

All authors listed above certify that they have NO affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

Acknowledgements

The authors would like to thank the Independent Hospital and Aged Care Pricing Authority (IHACPA) for enabling access to the relevant data and for providing technical support during the study, and Professor Stephen Duckett for advice and guidance.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.healthpol.2025.105416.

References

[1] ACSQHC. 'National safety and quality health service standards'. Australian Commission on Safety and Quality in Health Care; 2021. Accessed 8 December. https://www.safetyandquality.gov.au/standards/nsqhs-standards.
[2] ACSQHC. 'Hospital-acquired complications (HACs)'. Australian Commission on Safety and Quality in Health Care; 2022. Accessed 1 July, https://www.safetyandquality.gov.au/our-work/indicators/hospital-acquired-complications.
[3] ACSQHC. 'Healthcare-associated infection program'. Australian Commission on Safety and Quality in Health Care; 2024. Accessed 8 December, https://www.safetyandquality.gov.au/our-work/healthcare-associated-infection-program.
[4] AIHW. Australia's hospitals at a glance 2018-19. Canberra: Australian Institute of Health and Welfare; 2020.
[5] AIHW. Admitted patient care. Australian Insitute of Health and Welfare; 2024. Accessed 10 December, https://www.aihw.gov.au/reports-data/myhospitals/sectors/admitted-patients.
[6] AIHW. Admitted patient care NMDS 2024–25. Australian Institute of Health and Welfare; 2024. Accessed 10 October, https://meteor.aihw.gov.au/content/775630.
[7] AIHW. Australia's hospitals at a glance. Canberra: Australian Institute of Health and Welfare; 2024.
[8] AIHW. COVID-19. Canberra: Australian Institute of Health and Welfare; 2024.
[9] Barber Michael. How to run a government so that citizens benefit and taxpayers don't go crazy. London: Penguin Press; 2016.
[10] Bernal JL, Cummins S, Gasparrini A. 'Interrupted time series regression for the evaluation of public health interventions: a tutorial'. Int J Epidemiol 2017;46:348–55.
[11] Boudon Raymond. 'Limitations of rational choice theory'. Am J Sociol 1998;104:817–28.
[12] Centers for Medicare and Medicaid Services. 'Medicare program; changes to the hospital inpatient prospective payment systems and fiscal year 2008 rates'. Fed Regist 2007;72:47129–8175.
[13] Centers for Medicare and Medicaid Services. 'Hospital-acquired conditions (present on admission indicator)'. CMS; 2021. Accessed 12 July, https://www.cms.gov/medicare/medicare-fee-for-service-payment/hospitalacqcond.
[14] Charlson Mary, Szatrowski Ted P, Peterson Janey, Gold Jeffrey. 'Validation of a combined comorbidity index'. J Clin Epidemiol 1994;47:1245–51.
[15] COAG. "National health reform agreement - 2017 addendum.". Canberra: Council of Australian Governments; 2017.
[16] Duckett S, Jorm C, Moran G, Parsonage H. "Safer care saves money: how to improve patient care and save public money at the same time.". Melbourne: Grattan Institute; 2018.
[17] Duke GJ, Moran JL, Bersten AD, Bihari S, Roodenburg O, Karnon J, Hirth S, Hakendorf P, Santamaria JD. Hospital-acquired complications: the relative importance of hospital- and patient-related factors. Med J Aust 2022. https://doi.org/10.5694/mja2.51375.
[18] Edwards MJ, Duke GJ, Hakendorf P, Verghese S, Bihari S. 'Hospital acquired complications in South Australian major public hospitals'. Anaesth Intensive Care 2022;50:468–75.
[19] Ehsani JP, Jackson T, Duckett SJ. 'The incidence and cost of adverse events in Victorian hospitals 2003-04′. Med J Aust 2006;184:551–5.
[20] Health Policy Analysis. "Analysis of hospital-acquired diagnoses and their effect on case complexity and resource use – final report.". Sydney: Australian Commission on Safety and Quality in Health Care; 2013.
[21] IHACPA. "National pricing model technical specifications 2018–19: pricing and funding for safety and quality—risk adjustment model for hospital acquired complications.". Independent Hospital and Aged Care Pricing Authority; 2018.
[22] IHACPA. "IHACPA covid-19 response guidelines.". a. Independent Hospital and Aged Care Pricing Authority; 2020.
[23] IHACPA. "National efficient price determination 2020–21.". b. Sydney: Independent Hospital Pricing Authority; 2020.
[24] IHACPA. 'National benchmarking portal'. a. Independent Hospital and Aged Care Pricing Authority; 2024. Accessed 12 October, https://www.ihacpa.gov.au/health-care/data/national-benchmarking-portal.
[25] IHACPA. 'Safety and quality'. b. Independent Hospital and Aged Care Pricing Authority; 2024. Accessed 12 December, https://www.ihpa.gov.au/what-we-do/safety-and-quality.
[26] Kontopantelis E, Doran T, Springate DA, Buchan I, Reeves D. Regression based quasi-experimental approach when randomisation is not an option: interrupted time series analysis. BMJ 2015;350:h2750. https://doi.org/10.1136/bmj.h2750.
[27] McGuinness SL, Josphin J, Eades O, et al. Organizational responses to the COVID-19 pandemic in Victoria, Australia: a qualitative study across four healthcare settings. Front Public Health 2022;10:965664.
[28] McGuinness Sarah L, Johnson Josphin, Eades Owen, et al. Mental health outcomes in Australian healthcare and aged-care workers during the second year of the COVID-19 pandemic. Int J Env Res Public Health 2022;19:4951.
[29] McNair PD, Luft HS, Bindman AB. Medicare's policy not to pay for treating hospital-acquired conditions: the impact. Health Aff 2009;28:1485–93.
[30] NHFB. 'Funding reports'. National Health Funding Body; 2024. Accessed 12 October. https://www.publichospitalfunding.gov.au/public-hospital-funding-reports.
[31] Panagioti M, Khan K, Keers RN, et al. Prevalence, severity, and nature of preventable patient harm across medical care settings: systematic review and meta-analysis. BMJ 2019;366:l4185.
[32] Quan Hude, Vijaya Sundararajan, Patricia Halfon, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Med Care 2005;43:1130–9.
[33] R Core Team. "R: a language and environment for statistical computing.". Vienna: R Foundation for Statistical Computing; 2022.
[34] Ruggeri Kai, Alí Sonia, Berge Mari Louise, et al. Replicating patterns of prospect theory for decision under risk. Nat Hum Behav 2020;4:622–33.
[35] Scott S, Varian H. Predicting the present with Bayesian structural time series. Int J Math Model Numer Optim 2014;5(1/2). https://doi.org/10.1504/ijmmno.2014.059942.
[36] Shepheard Jennie. Clinical coding and the quality and integrity of health data. Health Inf Manag J 2020;49:3–4.
[37] Slawomirski L, Hensher M, Campbell J, deGraaff B. Pay-for-performance and patient safety in acute care: a systematic review. Health Policy 2024;143 (no pagination).
[38] Slawomirski Luke, Klazinga Nicolaas S. The economics of patient safety: from analysis to action. OECD health working papers. Paris: OECD; 2022.
[39] Tedesco D, Moghavem N, Weng Y, Fantini MP, Hernandez-Boussard T. 'Improvement in patient safety may precede policy changes: trends in patient safety indicators in the United States, 2000-2013′. J Patient Saf 2021;17:e327–ee34.
[40] Turner Simon L, Karahalios Amalia, Forbes Andrew B, Taljaard Monica, Grimshaw Jeremy M, McKenzie Joanne E. Comparison of six statistical methods for interrupted time series studies: empirical evaluation of 190 published series. BMC Med Res Methodol 2021;21:134.
[41] von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 2007;370:1453–7.
[42] Wagner AK, Soumerai SB, Zhang F, Ross-Degnan D. Segmented regression analysis of interrupted time series studies in medication use research. J Clin Pharm Ther 2002;27:299–309. https://doi.org/10.1046/j.1365-2710.2002.00430.x.
[43] WHO. "Global patient safety report 2024.". Geneva: World Health Organization; 2024.