

Approach To Map Health Cost Drivers

OMSAI THATI

omsaithati@gmail.com

JNTUH <https://orcid.org/0009-0001-4687-7225>

Method Article

Keywords: Universal Health Coverage Cube, Episode of Care Window, Bottom-Up Unit Micro-Costing, Generalized Linear Modelling, Health Cost Drivers

Posted Date: October 7th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11271811/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: The authors declare no competing interests.

Abstract

The WHO universal health coverage cube separates population, service, and cost coverage, yet at the point of care the three arrive as one episode: an entitlement, a service, and a claim met from public funds. That episode needs a unit; hospital cost studies each carry one dimension, in different cost units; this paper proposes the episode-of-care window, with a prespecified index episode and accumulation window, as the unit of analysis. Within the window, each dimension is operationalised as a vector: Entitlement for population (eleven variables, from age to days in hospital), Service Delivery for service (six cost categories, kept separate), and Reimbursement for cost coverage (insurance status, a binary non-cost variable). Because the dimensions arrive together, the three vectors enter one Gamma GLM with a log link to Direct Medical Costs, the episode's total cost. The exponentiated coefficients form the health cost driver map: the weight each variable carries within one episode. Applied to 125 households' episodes, procedure costs carry the largest z , with pharmacy, ICU, test and laboratory costs, disease type, prior disease, days in hospital, and discharge type significant; consultation fee and insurance are not. A separate robust run on the Entitlement and Reimbursement vectors is reported. The map is offered to ministries, insurers, and international bodies as a shared place to stand, informing their conversation without settling it, with each country bringing its own content. Held together at the point of care, the three dimensions of the UHC cube yield one map each institution can draw.

Introduction

Universal health coverage asks three questions at once: who is covered, which services they can use, and how much of the cost is protected from falling on the patient. The WHO UHC cube gives these questions shape through its three dimensions: population coverage, service coverage, and cost coverage. The World Health Organization and the World Bank's global monitoring reports treat service coverage and cost coverage as distinct dimensions [1]. This shared language connects the WHO, international health institutions, country ministries, and the organisations that work beside them. At the national level, it is well established. At the hospital, the three questions arrive as one episode. A person with a particular entitlement receives a particular set of services, and a bill is produced, part of which is met from public funds. This paper presents an approach to hold the three dimensions together at the point of care and read what drives the episode's cost as one map. It extends the cube, in the spirit of contribution, for decision-makers who want to know what they can do with the evidence and under what conditions. At the point of care, the three dimensions arrive as one episode, and each existing hospital cost study provides a complementary piece. Hospital cost analyses each address one part of the cost-allocation question, and each does so well. Hospital-level and patient-level studies identify demographic, clinical, utilisation, or institutional correlates of cost. Episode-costing approaches define service bundles and observation windows. Reimbursement studies examine payment incentives and cost recovery. These emphases appear in different designs: patient and visit expenditure in a Singapore academic medical centre, itemised ICU utilisation costing, hospital-charge predictors, episode-based cost assignment, and reimbursement and cost-recovery analyses [2, 3, 4, 5, 6]. Read together, the studies are

complementary. They carry patient attributes, service-activity quantities, and reimbursement rates, which are the same things the cube names as population, service, and cost [2, 3, 6, 7].

These studies also measure cost in different units. A prospective ICU study measured cost per ICU episode over a six-month enrolment period. A Medicare colonoscopy measure assigned clinically related services in a 14-day post-procedure window. The Singapore hospital study distinguished expenditure accumulated per patient over observed follow-up from expenditure per inpatient visit. These are not interchangeable estimands [2, 5, 8]. Each choice suits its purpose. Yet a ministry or an insurance agency that wished to combine such figures would be adding quantities that describe different things. A common unit has to be agreed before three dimensions can be held together. This extension defines the point of care as the episode-of-care window, and every cost and non-cost variable recorded within one episode is analysed together. With the episode-of-care window as the common unit, the extension is specified in one model, with precedent in the reviewed papers. The reviewed papers use Gamma GLM and log-link models, Gamma generalised estimating equations, multilevel Gamma models, linear or quantile regressions, sparse group lasso, predictive models, and episode-based cost-measure construction. The applied cost-regression evidence provides the Gamma family with a log link as precedent for the extension [7, 9, 10, 11, 12]. The extension takes the Gamma GLM with a log link and robust standard errors as its specification. It reads the exponentiated coefficients and the partial deviance explained as the health cost driver map. The map this model produces is offered in line with existing international practice. The framing also follows international practice on benefit packages. Benefit-package design considers whether services can be translated into delivery arrangements, resource requirements, and financing [13, 14]. The point-of-care health cost driver map extends this framing by considering entitlement, service activity, and reimbursement together within an episode of care. Within the hospital-costing and health-cost-modelling papers assembled for this study, the extension places the three vectors in one total-cost Gamma GLM with a log link and robust standard errors, scoped to this reviewed set [2, 3, 6, 7]. This matters beyond the hospital because countries often make coverage decisions with information produced for different purposes. A common unit and a common set of dimensions do not remove differences between countries. This assembled architecture is proposed as an “Approach To Map Health Cost Drivers” (ATMHCD).

Methods

The unit of analysis is the episode-of-care window, the point of care where the UHC cube meets the hospital variables, with a prespecified index episode and cost-accumulation window. Claims-based episode methods organise related services into an episode so that care and cost can be measured together [15, 16]. The stated boundaries are applied consistently to total cost and to every service-activity variable, so that nothing counted in one window is compared with something counted in another. The words encounter and episode refer to the same thing. Within this one window, the total cost of the episode is recorded alongside every service-activity variable. The dependent variable is Direct Medical Costs, the total patient hospital bill amount generated at the point-of-care encounter, which is the total cost of the episode. Billed charges, allowed amounts, and reimbursed amounts are not treated as

equivalent because they answer different questions. The reviewed set includes charge-as-cost proxy analysis, cost per ICU episode, per-visit expenditure, accumulated patient expenditure, and episode-based claims costing, which is why the definition is stated explicitly [2, 4, 8, 17].

The population dimension becomes the Entitlement vector. It contains Age, Gender, Region, Socioeconomic Status, Disease Type, Prior Disease, Disease Severity, Comorbid, Discharge Type, Ward Type, and Days in Hospital. Age is entered as a continuous variable, with its functional form and any spline or cohort sensitivity analysis stated. Clinical case-mix and socioeconomic or entitlement covariates are added if justified by the estimand. The Singapore hospital analysis included age, gender, ethnicity, a housing-based socioeconomic proxy, and resident status. The US Veterans Health Administration model used age, sex, and diagnostic categories [2, 18]. Alongside the Entitlement vector, the service dimension of the cube enters as the Service Delivery vector. Each service category enters as its own cost variable: Consultation Fee, Lab Costs, Test Costs, Procedure Costs, ICU Service Utilization, Pharmacy Costs; all recorded within the same episode window. Each category stays separate and is not collapsed into a single score. Bottom-up ICU costing has recorded resource use and reported laboratory, medication, and consumable cost outcomes. An emergency-admission study considered investigations and allied-professional time, and hospital-charge analysis highlighted length of stay and procedure counts [3, 4, 12]. Keeping categories separate preserves the granularity a planner needs, because a benefit package is a set of services, each with its own resource needs and financing. Alongside these two, the cost coverage dimension of the cube enters as the Reimbursement vector, held by Medical Insurance. It is the insurance status recorded for the encounter, coded binary at the episode level as insured or not insured. It enters the model as a non-cost variable, taken as recorded and read alongside Entitlement and Service Delivery as one more driver of the episode's total cost. Being insured is a fact about the episode, not a quantity computed from its cost, and its weight is read from the same exponentiated coefficient as the other variables. The Singapore study included ward subsidy categories and residency, the Washington State hospital study examined government reimbursement rates and provider costs, and the Indonesian stroke study analysed costs, reimbursement fees, and cost-recovery rate [2, 6, 19]. Each contributes a piece of the reimbursement vector. The available summaries do not establish a model in which this variable is entered jointly with a separate cost variable for each service category. The extension enters it jointly with the cost of each of the six service categories. Entered this way, insurance status, the marker of cost coverage in this article, sits on the same map as the other two dimensions, held together.

This model fits a prespecified Gamma GLM with a log link to episode total cost, entering all three vectors in the same model. Standard errors are robust and clustered at the level that matches repeated observations or shared providers, for example, patient or hospital. The variance estimator, the clustering unit, the finite-sample correction, and whether the robust variance is heteroskedasticity-consistent or cluster-robust are stated with the model. The combination is informed by applied Gamma/log-link and sandwich-GEE methods in the reviewed set [2, 7]. The one model, holding all three vectors, is then read as specified. The Gamma/log-link model is read with its link and variance assumptions, and its predicted and observed costs are stated [9, 20]. When observations are clustered, that structure is accounted for in

estimation, and comparative work has evaluated Gamma models under clustering and heteroscedasticity [21]. The prespecified Gamma model remains the primary specification, with its fit reported. Alongside the primary specification, robust standard errors and robust Gamma regression are reported separately. In simulation and cardiovascular-trial data, robust Gamma procedures were compared with the standard Gamma GLM. Results differed under contaminated data and for some covariate significance estimates [22]. For clustered cost data with heteroscedasticity, simulations compared Gamma generalised linear mixed models with log-linear mixed models, with performance depending on the data-generating conditions [21]. These comparisons support describing the variance estimator and clustering structure explicitly, so that the map is read with both stated.

The exponentiated coefficients, with the partial deviance explained, give the relative weight of each variable in the health cost driver map. The map is defined this way each time it appears, and it is always drawn from the one primary model. Staged comparisons are prespecified for reading the map: Entitlement plus Service Delivery plus Reimbursement; and the fully adjusted model. These comparisons sit around the primary unified fit and do not replace the simultaneous inclusion of the three vectors, which is the extension. The approach maps dissertation evidence grounded in 125 household episodes experienced at the point of care in advanced cardiovascular services across Suryapet District, as part of the Rajiv Aarogyasri sub-national safety net's policy transition toward Universal Health Coverage.

Results

The unit of analysis is the episode-of-care window, with a prespecified index episode and cost-accumulation window. The stated boundaries apply consistently to total cost and to every service-activity variable, and encounter and episode refer to the same thing. The dependent variable is Direct Medical Costs, the episode's total cost. The Entitlement vector contains Age, Gender, Region, Socioeconomic Status, Disease Type, Prior Disease, Disease Severity, Comorbid, Discharge Type, Ward Type, and Days in Hospital. The Service Delivery vector includes Consultation Fee, Lab Costs, Test Costs, Procedures Costs, ICU Service Utilization, and Pharmacy Costs. The Reimbursement vector is held by Medical Insurance.

Two models are run. The primary model fits a Gamma GLM with a log link to Direct Medical Costs, with all three vectors in the same model, OIM standard errors, $n = 125$. Its map is drawn from this one run. Alongside it, a separate Gamma GLM with a log link and robust standard errors is run on the non-cost variables only, the Entitlement and Reimbursement vectors, with eform, $n = 125$. Both models are run with their respective link and variance assumptions. The prespecified Gamma model remains the primary specification, with its fit reported. The robust run is reported separately alongside it.

The exponentiated coefficients give the relative weight of each variable in the health cost driver map, drawn from the one primary model. In the Entitlement vector, Disease Type carries 1.0712 ($z = 7.22$, $p < 0.001$), Prior Disease 0.9088 ($z = -3.98$, $p < 0.001$), Days in Hospital 1.0858 ($z = 3.83$, $p < 0.001$) and Discharge Type 1.0573 ($z = 2.08$, $p = 0.038$); the other Entitlement variables are not significant. In the

Service Delivery vector, Procedures Costs carries 1.0000043 ($z = 18.87$, the largest z in the model), Pharmacy Costs 1.0000107 ($z = 6.45$), ICU Service Utilization 1.0000067 ($z = 5.74$), Tests Costs 1.0000185 ($z = 2.55$) and Labs Costs 1.0000041 ($z = 2.32$), $p \leq 0.020$; Consultation Fee is not significant ($p = 0.429$). In the Reimbursement vector, Medical Insurance 1.0357 is not significant ($p = 0.157$). In the separate robust run on non-cost variables, Gender 0.7963 ($z = -3.08$, $p = 0.002$), Disease categories 1–4 ($p \leq 0.001$), Disease Severity level 2 (1.1834, $z = 2.28$, $p = 0.023$), Days in Hospital categories 1–4 ($p \leq 0.022$) and Ward 1.2378 ($z = 2.49$, $p = 0.013$) are significant; Age, Region, Insurance, Comorbid and Discharge Type are not.

Table 1
Health Cost Driver Map – Primary Model Gamma GLM, Log Link, OIM Standard Errors, $n = 125$

Vector	Variable	$\exp(\beta)$	z	p
Entitlement	Age	1.0008833	1.20	0.229
	Gender	0.9880033	-0.35	0.727
	Region	1.0173732	0.74	0.458
	Socioeconomic Status	1.0172377	0.71	0.476
	Disease Type	1.0712659	7.22	< 0.001
	Disease Severity	1.0040820	0.24	0.811
	Comorbid	0.9664224	-0.91	0.364
	Prior Disease	0.9088160	-3.98	< 0.001
	Days in Hospital	1.0857376	3.83	< 0.001
	Ward Type	1.0111119	0.35	0.725
	Discharge Type	1.0572439	2.08	0.038
Reimbursement	Medical Insurance	1.0356768	1.42	0.157
Service Delivery	Consultation Fee	1.0000973	0.79	0.429
	Labs Costs	1.0000041	2.32	0.020
	Tests Costs	1.0000185	2.55	0.011
	Pharmacy Costs	1.0000107	6.45	< 0.001
	Procedures Costs	1.0000043	18.87	< 0.001
	ICU Service Utilization	1.0000067	5.74	< 0.001

Deviance/df = 0.0109 · AIC = 25.72 · BIC = – 510.65 · Constant = 10.1402

Table 2
Separate Robust Run – Non-Cost Variables Gamma GLM, Log Link, Robust
 Standard Errors, eform, n = 125

Vector	Variable	exp(β)	z	p
Entitlement	Age (= 1)	0.9662535	–0.77	0.442
	Gender (= 1)	0.7962948	–3.08	0.002
	Region (= 1)	1.0249660	0.46	0.648
	SES (= 1)	0.8788867	–1.78	0.075
	SES (= 2)	0.9133134	–1.10	0.272
	Disease (= 1)	1.3828280	3.27	0.001
	Disease (= 2)	1.4113580	3.38	0.001
	Disease (= 3)	1.5788510	3.33	0.001
	Disease (= 4)	1.6980040	5.13	< 0.001
	Severity (= 1)	0.9707548	–0.48	0.629
	Severity (= 2)	1.1833740	2.28	0.023
	Comorbid (= 1)	0.9164409	–1.17	0.241
	Prior Disease (= 1)	0.9601421	–0.71	0.478
	Days in Hospital (= 1)	1.3521850	2.29	0.022
	Days in Hospital (= 2)	1.8063340	4.06	< 0.001
	Days in Hospital (= 3)	3.0796130	7.61	< 0.001
	Days in Hospital (= 4)	3.5944420	6.22	< 0.001
	Ward (= 1)	1.2377680	2.49	0.013
	Discharge (= 1)	1.0020620	0.03	0.975
Reimbursement	Insurance (= 1)	1.0656650	1.28	0.201

Deviance/df = 0.0483 · AIC = 26.19 · BIC = – 497.13 · Constant = 62603.27

Discussion

The existing papers provide complementary pieces, and the extension assembles the pieces into one. Hospital cost-driver studies establish the relevance of patient and clinical attributes, length of stay, procedures, and resource use. Hospital charges were associated with diagnostic category, length of stay, procedure counts, operating-room procedures and hospital ownership, while emergency-admission episode costs were strongly related to clinical complexity and resource use [4, 12, 23]. Bottom-up ICU costing itemised laboratory tests, medication and consumables, but modelled costs with linear regression [3]. Episode-costing work shows that cost attribution depends on episode boundaries and on which services are assigned to the episode [5, 24]. Reimbursement appears in a separate cluster of evidence. Washington State hospital panel data examine changes in patient mix when Medicare and Medicaid reduce reimbursement rates. An Indonesian hospital study analyses patient cost, reimbursement fees, and cost-recovery rate [6, 19]. Their available summaries do not establish a common patient-level model that simultaneously contains entitlement variables, separate service-category costs, and insurance status. This extension places the entitlement variables, the six service cost categories, and insurance status in one point-of-care run. The extension holds these complementary pieces in one point-of-care run, and the Gamma-family studies give it precedent. Gamma-family approaches are represented by several methods and applications. The Singapore paper reports Gamma/log-link models and sandwich variance for visit-level GEE. The palliative-care cost paper reports Gamma/log-link GLMs, the multicountry stroke paper compares multilevel Gamma and normal models, and methodological papers compare Gamma regression or robust Gamma alternatives. These studies differ in outcome, hierarchy, covariance structure, and purpose. They are the methodological precedent for the extension and not a direct replication of its unified model [2, 7, 10, 21, 22, 25, 26]. The novelty is the explicit co-specification and evaluation of the three vectors in this dataset and model, and the papers illustrate complementary, non-identical models across hospital costing, Gamma methods, reimbursement, and episode construction [2, 5, 19, 22]. Among these precedents, the Singapore hospital study is the closest precedent to the extension. It separates patient-level and visit-level expenditure models, reports Gamma/log-link models, and uses sandwich variance estimation for repeated-visit GEE. It includes residency and ward subsidy class, while excluding utilisation measures such as length of stay and ICU days from expenditure predictors. The available report does not show separate service-category costs alongside entitlement variables and insurance status in that same model. The extension adds both to the same model [2]. No claim is made against that work. It is the nearest piece, and the extension builds on it. Built on this nearest piece, the extension turns to what decision-makers can do with the evidence. For institutions that support countries, the value of the approach lies in the options it opens, offered with their conditions and not as verdicts. A ministry of health designing or revising a benefit package could use a point-of-care map to see which service categories carry the most weight in an episode. It could then consider whether delivery arrangements or resource planning in those categories deserve attention, in line with the attention that benefit-package work gives to delivery and financing [13, 14]. A national health insurance agency could look at how insurance status sits beside the service variables in the same model, and consider what that suggests for its own settings, recognising that its payment rules, population and data will shape what the map shows. An international organisation working with several countries could use the common unit and the three common vectors as a shared

vocabulary, so that discussions between ministries rest on comparable definitions. In each case, the map informs a conversation and does not settle it. The conclusions belong to those who hold the responsibility for them.

The fit of the primary model is Deviance/df = 0.0109, AIC = 25.72, and BIC = - 510.65 at n = 125. The fit of the separate robust run is Deviance/df = 0.0483, AIC = 26.19, and BIC = - 497.13. Under the Gamma family with a log link, the lower deviance per degree of freedom in the primary model reflects the cost variables; their inclusion brings the unexplained cost variation closer to the predicted cost. The deviance belongs to each model as a whole and does not divide the weight among the vectors, which the exponentiated coefficients do, variable by variable. This is consistent with how UHC is monitored internationally, where service coverage and cost coverage are treated as distinct dimensions [1]. The point-of-care map extends the same framing downward, to entitlement, service activity and reimbursement within one episode of care. Countries differ in how they define entitlement, organise services and finance care. An approach that holds the three dimensions in one place without prescribing any arrangement leaves room for each country to see its own situation in the map. It offers a common structure, and each country brings its own content.

Conclusion

This paper acknowledges the studies that informed it and recognises the shared language of universal health coverage. Among the papers in this reviewed set, no paper explicitly documents all three specified vectors together in a single total-cost Gamma GLM with a log link and robust standard errors [2, 3, 5, 6, 7]. The extension offers the episode-of-care window as a common unit, and the Entitlement, Service Delivery and Reimbursement vectors as a shared frame for the three dimensions of the WHO UHC cube. It offers each institution a health cost driver map of its own. The extension takes no position on how any country defines entitlement, organises services or finances care; each country brings its own content to this frame. The authors invite ministries, agencies and partner organisations to apply the extension to their own data and settings, and to begin drawing their own map from the first episode. The authors welcome further material from other institutions that was not available to this review. The extension is offered in a spirit of contribution, in the hope that the language of universal health coverage will come closer to where care is delivered and paid for.

Declarations

Acknowledgements section

The authors thank Adepu Ramesh for supervising.

Declaration of conflicting interest

The authors declare no conflicting interests.

Funding statement

This work received no external funding.

Ethical approval

The study was approved by the Institutional Human Ethics Committee of Vikas College of Pharmaceutical Sciences, Suryapet, affiliated with JNTUH, India (approval no. VCPS/IHEC/2024-2025/002), in accordance with the Declaration of Helsinki.

Informed consent statement

Written informed consent for participation and publication was obtained from all individual participants included in the study.

Data availability statement

The datasets are confidential due to patient privacy; however, the processed dataset can be shared upon serious consideration.

Use of artificial intelligence

Claude assisted with wording and clarity; the authors reviewed and approved all content.

Authors contributions

Omsai Netha Thati: Conceptualization, Methodology, Formal analysis, Investigation, Writing – original draft, Writing – review and editing. Guntur Divya, Velugupally Godasri, Rumana: Methodology, Investigation, Writing – review and editing. All authors approved the final manuscript and accept full responsibility for the study.

References

1. World Health Organization, International Bank for Reconstruction and Development/The World Bank (2025) Tracking universal health coverage: 2025 global monitoring report. World Health Organization, Geneva
2. Rahman N, Ng SH, Ramachandran S, Wang DD, Sridharan S, Tan CS et al (2019) Drivers of hospital expenditure and length of stay in an academic medical centre: a retrospective cross-sectional study. BMC Health Serv Res 19(1):442

3. Mastrogianni M, Galanis P, Siskou O, Kaitelidou D, Katsoulas T (2023) Development of a variable cost prediction model for hospitalized critically ill patients in ICU. *Health Rev* 34:12–17
4. Aplin B, Nazzal M, Qu W, Zelenock G, Kazan V, Abbas J (2015) Patient variables impacting hospital costs from 2008 to 2010. *Am J Surg* 210(1):134–140
5. Park WG, Sandhu A, MaCurdy T (2021) Development of a cost measure for screening/surveillance colonoscopy for the Merit-Based Incentive Payment System. *Gastroenterology* 161(1):4–7e4
6. Fahlevi H, Putra TRI, Oktari RS (2020) Predicting cost recovery rate of ischemic stroke patients: a potential application of big data analysis in hospital. *J Phys Conf Ser* 1500(1):012106
7. May P, Garrido MM, Cassel JB, Morrison RS, Normand C (2016) Using length of stay to control for unobserved heterogeneity when estimating treatment effect on hospital costs with observational data: issues of reliability, robustness, and usefulness. *Health Serv Res* 51(5):2020–2043
8. Moran JL, Solomon PJ, Peisach AR, Martin J (2007) New models for old questions: generalized linear models for cost prediction. *J Eval Clin Pract* 13(3):381–389
9. Dodd S, Bassi A, Bodger K, Williamson P (2006) A comparison of multivariable regression models to analyse cost data. *J Eval Clin Pract* 12(1):76–86
10. Grieve R, Nixon R, Thompson SG, Normand C (2005) Using multilevel models for assessing the variability of multinational resource use and cost data. *Health Econ* 14(2):185–196
11. Swierkowski P, Barnett A (2018) Identification of hospital cost drivers using sparse group lasso. *PLoS ONE* 13(10):e0204300
12. Cournane S, Byrne D, O'Riordan D, Fitzgerald B, Silke B (2016) The relationship between resource utilization, clinical risk and hospital costs for emergency medical admissions. *Acute Med* 15(1):7–12
13. Reynolds T, Wilkinson T, Bertram MY, Jowett M, Baltussen R, Mataria A et al (2023) Building implementable packages for universal health coverage. *BMJ Glob Health* 8(Suppl 1):e010807
14. El-Jardali F, Fadlallah R, Daouk A, Rizk R, Hemadi N, El Kebbi O et al (2019) Barriers and facilitators to implementation of essential health benefits package within primary health care settings in low-income and middle-income countries: a systematic review. *Int J Health Plann Manage* 34(1):15–41
15. Schulman KA, Yabroff KR, Kong J, Gold KF, Rubenstein LE, Epstein AJ et al (1999) A claims data approach to defining an episode of care. *Health Serv Res* 34(2):603–621
16. Hornbrook MC, Hurtado AV, Johnson RE (1985) Health care episodes: definition, measurement and use. *Med Care Rev* 42(2):163–218
17. Huang S, Hartman B, Brazauskas V (2017) Model selection and averaging of health costs in episode treatment groups. *ASTIN Bull* 47(1):153–167
18. Montez-Rath M, Christiansen CL, Ettner SL, Loveland S, Rosen AK (2006) Performance of statistical models to predict mental health and substance abuse cost. *BMC Med Res Methodol* 6:53
19. Friesner D, Rosenman R, McPherson M (2005) The effects of government reimbursement on hospital costs: some empirical evidence from Washington State. *Mt Plains J Bus Technol.* ;6(1).

20. Zhou J, Williams C, Keng MJ, Wu R, Mihaylova B (2024) Estimating costs associated with disease model states using generalized linear models: a tutorial. *PharmacoEconomics* 42(3):261–273
21. Baldi I, Pagano E, Berchialla P, Desideri A, Ferrando A, Merletti F et al (2013) Modeling healthcare costs in simultaneous presence of asymmetry, heteroscedasticity and correlation. *J Appl Stat* 40(2):298–310
22. Petrinco M, Barbatì G, Meylan D, Pagano E, Gregori D, Merletti F et al (2012) Robust gamma regression models for the analysis of health care cost data. *Model Assist Stat Appl* 7(2):115–124
23. Carey K (2000) Hospital cost containment and length of stay: an econometric analysis. *South Econ J* 67:363–380
24. Geue C, Lewsey J, Lorgelly P, Govan L, Hart C, Briggs A (2012) Spoilt for choice: implications of using alternative methods of costing hospital episode statistics. *Health Econ* 21:1201–1216
25. Locatelli I, Marazzi A (2013) Robust parametric indirect estimates of the expected cost of a hospital stay with covariates and censored data. *Stat Med* 32(14):2457–2466
26. Thompson SG, Nixon RM, Grieve R (2006) Addressing the issues that arise in analysing multicentre cost data, with application to a multinational study. *J Health Econ* 25(6):1015–1028