Using the Johns Hopkins ACG Case-Mix System for population segmentation in a hospital-based adult patient population in Singapore

Joshua Kuan Tan ,1 Xiaojin Zhang,1 Dawn Cheng,2 Ian Yi Onn Leong,3 Chia Siong Wong,2 Jeannie Tey,4 Shu Ching Loh,3 Eugene Fidelis Soh,5 Wei Yen Lim1

ABSTRACT

Objective Population health management involves risk characterisation and patient segmentation. Almost all population segmentation tools require comprehensive health information spanning the full care continuum. We assessed the utility of applying the ACG System as a population risk segmentation tool using only hospital data.

Design Retrospective cohort study.

Setting Tertiary hospital in central Singapore.

Participants 100 000 randomly selected adult patients from 1 January to 31 December 2017.

Intervention Hospital encounters, diagnoses codes and medications prescribed to the participants were used as input data to the ACG System.

Primary and Secondary Outcome Measures Hospital costs, admission episodes and mortality of these patients in the subsequent year (2018) were used to assess the utility of ACG System outputs such as resource utilisation bands (RUBs) in stratifying patients and identifying high hospital care users.

Results Patients placed in higher RUBs had higher prospective (2018) healthcare costs, and were more likely to have healthcare costs in the top five percentile, to have three or more hospital admissions, and to die in the subsequent year. A combination of RUBs and ACG System generated rank probability of high healthcare costs, age and gender that had good discriminatory ability for all three outcomes, with area under the receiver-operator characteristic curve (AUC) values of 0.827, 0.889 and 0.876, respectively. Application of machine learning methods improved AUCs marginally by about 0.02 in predicting the top five percentile of healthcare costs and death in the subsequent year.

Conclusion A population stratification and risk prediction tool can be used to appropriately segment populations in a hospital patient population even with incomplete clinical data.

INTRODUCTION

Many countries around the world have experienced demographic and epidemiological transitions, with an increase in the population prevalence of chronic conditions such as diabetes, hypertension and hyperlipidaemia, and face escalating healthcare utilisation, which has strained healthcare resources.1 2 Population health management, with a focus on prevention and care integration, has been a key strategy to manage these resource constraints.3 A key component of population health management is population risk segmentation, which aims to stratify a heterogeneous population into relatively homogeneous and distinct subgroups based on their healthcare needs and prospective utilisation patterns.3–5 Population risk segmentation allows healthcare administrators to better understand their population, identify unmet healthcare needs within a population segment and address these through integrated care intervention programmes, supported by appropriate financing frameworks.4–6

Singapore has likewise experienced a very rapid demographic transition within the last 20 years where the proportion of persons 65 years and older has increased from 7.2% in 2000 to 15.2% in 2020,7 coupled with an epidemiological transition that started in the 1960s and was largely completed by the 1980s.8 These secular changes have strained the country’s healthcare system.9 10 Singapore’s
healthcare system comprises a network of private primary care clinics, public polyclinics, public hospitals, tertiary-specialist care centres and private hospitals. Care provision is fairly fragmented: the private sector, comprising hundreds of individual providers, is the main provider of primary care (about 80% of the market), while the public sector, comprising three large regional health systems, is the major provider of hospital care (also about 80% of that market). In addition, providers use different electronic (and in some cases, paper) systems, with limited sharing of patient medical information at this point.

A key recent development in Singapore has been the implementation of the Regional Health System (RHS) framework, in which the country is divided into three geographical regions and healthcare is coordinated within each region. Each RHS consists of at least two general hospitals working closely with other healthcare providers (primary care providers, community hospitals, nursing homes, home care and day rehabilitation providers) and social care providers to meet the needs of the population in that geographical region. In 2022, the Ministry of Health unveiled a new Population Health Management plan. To support population health management by our RHS, use of administrative data to perform population segmentation and identify high-risk patients is a key enabling tool. Tan Tock Seng Hospital (TTSH) hence has started exploring different approaches to developing a segmentation and risk identification framework.

The Johns Hopkins ACG System (ACG System) is a diagnosis-based, case-mix methodology that has been used by various healthcare organisations for population segmentation. The ACG System has been validated in the USA, Canada, several countries in Europe and in Asia, particularly in Taiwan, and is used to identify high-risk patients for interventions. Validation studies of the ACG System have largely taken place in the context of a single provider, using comprehensive health utilisation information available to that provider that spanned the full care continuum across primary, tertiary and step-down and community care, and as far as we know, it has not been performed in the context of a hospital service provider. In this study, we assessed whether the ACG System can be used for population risk segmentation among patients in a tertiary setting at the TTSH, a 1600-bed public hospital with specialist services serving a large catchment area in central Singapore, using only data available to the hospital as system input.

METHODS

ACG System

We used V.12.0 of the ACG System (released October 2019), which has been extensively described. In brief, the ACG System collapses diagnosis codes into Aggregated Diagnoses Groups (ADGs) on the basis of similarity in healthcare needs. Based on the number and type of assigned ADGs, each patient is then assigned to an Adjusted Clinical Group (ACG). AGCs are further categorised into five resource utilisation bands (RUBs 1–5; with higher RUBs indicating patients with greater morbidity requiring higher healthcare utilisation). The ACG System generates other patient markers such as frailty indicators, and probabilities for relevant outcomes in the subsequent year (including high total cost, inpatient admission, extended admissions, 30-day readmission, high pharmacy costs and predicted cost range). Apart from these, the ACG System also generates model markers for each patient that can be used to further calibrate risk prediction models. In total, the ACG System generated 617 model markers.

Data collection

Our study population was a random sample of 100 000 patients who had contact with TTSH either through the emergency department, hospital admission, day surgery centre, medical day centre or outpatient settings (Specialist Outpatient Clinics (SOCs) and Health Enrichment Centre) between 1 January 2017 and 31 December 2017, from a total population of 369 267 patients who had at least one documented contact with the hospital.

Data of selected patients were extracted from electronic medical records and hospital financial records, and included age, gender, all diagnoses (International Classification of Disease (ICD)-9, ICD-10-Australian Modification (AM), Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT)), medications dispensed (including date of dispensing and number of refill days), pharmacy-specific cost information, medical service provider information, specific procedure and treatment service information (including dialysis, mechanical ventilation, cancer therapy and psychotherapy), utilisation information (including number of inpatient hospitalisation episodes, inpatient hospitalisation days, count of 30-day readmission, count of unplanned 30-day readmission, number of emergency visits, number of SOC visits) and total cost information for the period 1 January 2017–31 December 2018. Costs refer to hospital charges before any government subsidies. Only data captured within TTSH were available, and no medical or financial information from other healthcare providers (including primary care, community pharmacies, other hospital providers and step-down services) was available for this analysis.

Statistical analysis

Data from 1 January 2017 to 31 December 2017 were used as input data for the ACG System. We used ACG System outputs as independent variables to predict outcomes of interest: total healthcare costs in TTSH in the subsequent year (absolute costs and costs in the top 5%), three or more hospital admissions in the subsequent year and mortality in the subsequent year. Healthcare utilisation and mortality data from 1 January 2018 to 31 December
2018 were used to compare predicted outcomes with observed outcomes in 2018. We removed patients who had died by 31 December 2017. Linear and logistic regression models were used. The adjusted R² was used to measure the explanatory power for the linear regression. The area under the receiver-operator characteristic curve (AUC) was used to evaluate the discriminatory abilities of logistic models. The mean squared error was also reported as a measure of statistical fit.

We further explored whether the 617 model markers assigned by the ACG System to each patient in our sample could improve prediction of prospective healthcare utilisation. We divided our sample randomly (70:30) into training and testing sets, calibrated the models on the training set using variable selection methods (stepwise, forward and backward regression with least angle regression and least absolute shrinkage and selection operator regression). We also created sequential two-stage models: the first stage was a class model using logistic regression to predict the probability that a patient will return in the next year (ie, cost in 2018 ≠ 0); the second stage incorporated the predicted probability from the first stage into either a value (ie, linear regression) or class model (ie, logistic regression, neural network, random forest, decision tree). The best model was selected based on the Akaike Information Criteria or Schwartz Bayesian Criteria at a significance level of p<0.05. Adjusted R² or AUC and mean absolute error in the testing set were reported.

Data used in this study were prepared using SAS Enterprise Guide software V.6.1,30 and statistical analysis was performed using Stata V.14.031 and SAS Enterprise Miner software V.13.2.32

Patient and public involvement
No patient involved.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Study population age structure and gender distribution compared with Singapore's resident population in 2017</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Population gender distribution by age bands</strong></td>
<td><strong>Singapore's resident population in 2017</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Study population in 2017</strong></td>
</tr>
<tr>
<td><strong>Age band (years)</strong></td>
<td></td>
</tr>
<tr>
<td>00–04</td>
<td>0.10</td>
</tr>
<tr>
<td>05–11</td>
<td>0.25</td>
</tr>
<tr>
<td>12–17</td>
<td>0.51</td>
</tr>
<tr>
<td>18–34</td>
<td>9.86</td>
</tr>
<tr>
<td>35–44</td>
<td>5.90</td>
</tr>
<tr>
<td>45–54</td>
<td>6.90</td>
</tr>
<tr>
<td>55–64</td>
<td>10.05</td>
</tr>
<tr>
<td>65–69</td>
<td>5.26</td>
</tr>
<tr>
<td>70–74</td>
<td>4.08</td>
</tr>
<tr>
<td>75–79</td>
<td>3.14</td>
</tr>
<tr>
<td>80+</td>
<td>3.72</td>
</tr>
<tr>
<td><strong>Subtotal</strong></td>
<td>49.77</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
</tr>
</tbody>
</table>

RESULTS

The age and gender distribution of the study population is described in table 1 (also online supplemental figure 1). Across the age bands, the gender distribution was balanced. The age distribution of the study population was skewed towards older adults when compared with Singapore’s resident population in 2017.33 This reflects the fact that the sample was taken from a patient population of a general hospital without neonatal, paediatric, obstetrics and gynaecology departments, and older persons are more likely to present to hospitals. Overall, the study population was comparable with the TTSH patient population in 2017 (online supplemental table 1).

Total expenditure for these 100 000 patients in 2017 was 341 108 702.00 in Singapore dollars (SGD$), with an average expenditure of SGD$3411.08.00. By 31 December, 1943 patients had died, and these individuals were removed from subsequent analyses. In 2018, total expenditure was SGD$233 802 869.00, and 41 048 of these 100 000 patients did not have any encounter with TTSH. Among those with at least one encounter, average expenditure was SGD$3965.99. A total of 14.70% (n=14 744) of the population had at least one admission, among whom 1347 patients had three or more admissions in 2017.

The ACG System classified 4.37% of the study population in RUB 5 (‘very high risk’), and a further 3.65% of the population in RUB 4 (‘high risk’). The RUB 3 population was 24.55% of the total (figure 1). There was a larger proportion of patients in high-risk categories (RUBs 4 and 5) among older age bands, and this is uniform for both genders (figure 2).

The population in RUB 5, despite its small size, accounted for the largest proportion (37.58%) of cost
in 2017 (figure 1), but this decreased to 18.35% of cost in 2018. The mean cost per patient was SGD$542.78 for RUB 1, $1270.58 for RUB 2, SGD$4240.78 for RUB 3, SGD$13 507.00 for RUB 4 and SGD$29 304.09 for RUB 5 in 2017. Costs in 2018 likewise increased from RUB 1 to RUB 5 in 2018 (figure 3). However, there was wide variation in costs in both years.

We observed a stepwise increase in the proportion of persons with costs in top 5% in 2018, three or more hospital admissions in 2018, or died in 2018, across RUB categories (table 2). Almost one-third of patients in RUB 5 had costs in the top 5% in 2018, about 12% had three or more admissions in 2018, and 15% died in 2018.

Age and gender alone (model 1) accounted for about 14% of the variance in total costs accrued in 2018; the variance increased to 15% when length of stay (LOS) was included in the model (model 2), and models that included ACG markers on top of these two variables had a much greater $R^2$ than the age and gender model alone (table 3). A two-stage model that first predicted whether a patient would have an encounter in 2018, before predicting costs in 2018, selecting factors from among the 617 model markers using a machine learning approach, significantly improved prediction and explained about 38% of the variance in costs.

Models using age, gender and LOS gave fair-to-good predictive ability for subsequent year costs in the top five percentile, three or more hospital admissions in the subsequent year and death in the subsequent year. One-stage logistic models using ACG markers had fair-to-good discriminatory ability in predicting costs in the top five percentile, three or more hospital admissions and death in 2018, with AUCs in the 0.73–0.88 range (table 4). A model using age, gender and RUB gave better predictive ability compared with the model using age, gender and LOS for top five percentile of costs and three or more hospital admissions, but not mortality. The use of four predicting variables (RUB, the rank probability of high

Figure 1 Distribution of study population and total cost in 2017 and 2018 by resource utilisation bands (RUBs).

Figure 2 Distribution of resource utilisation band (RUB) by age groups and gender.
cost (RPHC), age and gender) gave very good discriminatory ability for these three outcomes. Machine learning approaches showed only a marginal improvement in AUC of about 0.02–0.03, when compared with standard logistic regression models.

**DISCUSSION**

Overall, this study suggests that outputs from the ACG System can be used to segment a hospital population by risk of subsequent year adverse outcomes and utilisation, even though only partial healthcare information about these patients is available. While a model using age, gender and LOS gave fair-to-good AUCs when predicting top 5% patients in cost, patients needing three or more admissions and mortality, a model using age, gender and RUB, gave a 0.06 (about 7%–8%) improvement in AUC for the first two outcomes, and comparable AUCs for mortality. Models using RUBs alone are meaningfully associated with prospective high costs, mortality and frequent admissions. Further, use of RUB in conjunction with age, gender and the RPHC gives good discriminatory ability in identifying patients with extremely high utilisation. The addition of ACG System model markers with a machine learning approach improved quite substantially the prediction of absolute costs (R² of 0.385 vs 0.154 for a simple model using age, gender and LOS).

Other validation studies have shown that the ACG System can be used for population segmentation and identification of high-risk individuals, in particular in the setting of enrolled populations within a managed health-care programme, and where health information about these populations is fairly complete. Our study adds to this knowledge by showing that the ACG System can also be used to segment hospital populations and identify high-risk individuals, even in a case where healthcare information available from that hospital population is incomplete. The adjusted R² that we obtained for predicting prospective costs and the AUCs for predicting prospective top 5% users are similar to those reported by studies that used community populations enrolled into managed health programmes, and where health data were relatively complete.

**Table 2** Counts (proportions) of patients within each RUB with the outcome of interest

<table>
<thead>
<tr>
<th>Outcome</th>
<th>RUB 1</th>
<th>RUB 2</th>
<th>RUB 3</th>
<th>RUB 4</th>
<th>RUB 5</th>
<th>P value (Χ² test)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top 5% cost in 2018</td>
<td>583 (1.63%)</td>
<td>811 (2.59%)</td>
<td>1888 (7.78%)</td>
<td>624 (18.58%)</td>
<td>1094 (32.17%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>At least 3 admissions in 2018</td>
<td>66 (0.19%)</td>
<td>121 (0.39%)</td>
<td>370 (1.53%)</td>
<td>173 (5.15%)</td>
<td>406 (11.94%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Died in 2018</td>
<td>299 (0.84%)</td>
<td>258 (0.82%)</td>
<td>480 (1.98%)</td>
<td>206 (6.13%)</td>
<td>511 (15.02%)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Analyses excluded patients who died in 2017.

RUB, resource utilisation band.

**Table 3** Performance of regression models in predicting prospective costs

<table>
<thead>
<tr>
<th>Model for predicting prospective (2018) costs</th>
<th>Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1: age and gender</td>
<td>Adjusted R²=0.147</td>
</tr>
<tr>
<td>Model 2: age, gender and length of stay in 2017</td>
<td>Adjusted R²=0.154</td>
</tr>
<tr>
<td>Model 3: rank probability of high cost (RPHC)</td>
<td>Adjusted R²=0.083</td>
</tr>
<tr>
<td>Model 4: age, gender, RPHC, RUB</td>
<td>Adjusted R²=0.236</td>
</tr>
<tr>
<td>Model 5: age, gender, RPHC, RUB, frailty</td>
<td>Adjusted R²=0.237</td>
</tr>
<tr>
<td>Model 6: two-stage model*</td>
<td>Adjusted R²=0.385; mean squared error=7.662</td>
</tr>
</tbody>
</table>

*Best model: stage 1—forward variable selection and AIC model selection criterion; stage 2—backward variable selection with AIC model selection criterion. AIC, Akaike Information Criteria; RUB, resource utilisation band.
The literature suggests that a data-driven approach to population segmentation is possible. The Guided Care Programme in the USA has successfully used predictive risk modelling to identify high-risk patients for enrolment into case management. In Spain, the Ribera Salud group used a population stratification model based on a variety of tools, including the ACG System, and implemented a Population Health Management Programme called ‘Trucare’. The National Health Service in England has promoted the use of a variety of population segmentation and risk identification tools to flag out high-risk patients for closer reviews and management.

For the clinician, it is essential that any segmentation strategy provides clear information about what the clinician can do for patients who have been identified to be at high risk. Realistically, the segmentation model may not provide significant additional information to the clinician, who has the patient in front of her and can obtain any further information that she needs from the patient, but a segmentation model that provides an alert to clinicians about potential high-risk patients can serve as a prompt to the clinician to perform further assessment of the patient, which may include assessment of social or financial needs beyond just the health needs.

Population segmentation may confer some advantages to population health management beyond individual-level patient care. Population segmentation and risk identification could be used to guide the frequency of follow-ups of patients and the nature of intervention (eg, multidisciplinary case management for patients at higher risk bands, and a lighter touch telephone review, or online application-based contact for lower-risk patients), as well as notify clinicians of patients who appear to be ‘progressing’ in risk. This segmentation system could be used for risk adjustment when evaluating the performance of different providers within the RHS, or for financial payments. Finally, risk identification could identify high-risk patients who have defaulted appointments or are otherwise not on follow-up, and allow these patients to be called back for assessment and management.

Even after deciding on a segmentation model, significant work still needs to be done so that segmentation can be applied effectively as a strategy to improve population health. This will include developing the information technology solutions to efficiently segment patients based on a population segmentation model, developing the feedback mechanisms to provide segmentation results to clinicians, operations managers and population health managers, and setting up feedback mechanisms to provide segmentation results to clinicians, operations managers and population health managers.

### Table 4: Prediction models for top 5% of prospective healthcare costs, three or more hospital admissions in 2018 and death in 2018

<table>
<thead>
<tr>
<th>Model</th>
<th>Predicting costs in top 5 percentile in 2018</th>
<th>Predicting 3 or more hospital admissions in 2018</th>
<th>Predicting death in 2018</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1: age and gender</td>
<td>AUC=0.731, MSE=0.047</td>
<td>AUC=0.748, MSE=0.011</td>
<td>AUC=0.828, MSE=0.017</td>
</tr>
<tr>
<td>Model 2: age, gender and LOS in 2017</td>
<td>AUC=0.766, MSE=0.046</td>
<td>AUC=0.802, MSE=0.011</td>
<td>AUC=0.874, MSE=0.016</td>
</tr>
<tr>
<td>Model 3: RUB</td>
<td>AUC=0.758, MSE=0.044</td>
<td>AUC=0.748, MSE=0.011</td>
<td>AUC=0.732, MSE=0.016</td>
</tr>
<tr>
<td>Model 4: RUB, age, gender</td>
<td>AUC=0.827, MSE=0.043</td>
<td>AUC=0.877, MSE=0.011</td>
<td>AUC=0.867, MSE=0.017</td>
</tr>
<tr>
<td>Model 5: RPHC</td>
<td>AUC=0.802, MSE=0.044</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 6: RPHC, age, gender</td>
<td>AUC=0.815, MSE=0.044</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 7: RPHC, RUB, age, gender</td>
<td>AUC=0.785, MSE=0.046</td>
<td>AUC=0.889, MSE=0.011</td>
<td>AUC=0.876, MSE=0.016</td>
</tr>
<tr>
<td>Model 8: one-stage machine learning model* (best model: forward selection with significant level stop criterion)</td>
<td>AUC=0.856, MSE=0.040</td>
<td>AUC=0.842, MSE=0.011</td>
<td>AUC=0.894, MSE=0.015</td>
</tr>
<tr>
<td>Model 9: two-stage machine learning model† (best model: stages 1 and 2—backward variable selection with AIC model selection criterion)</td>
<td>AUC=0.856, MSE=0.040</td>
<td>AUC=0.842, MSE=0.011</td>
<td>–</td>
</tr>
</tbody>
</table>

*Best model: for costs—forward variable selection with significant level stop criterion; for admissions—forward variable selection with SBC model selection criterion; for death—stepwise selection with AIC model selection criterion.
†Best model: for costs (stages 1 and 2)—backward variable selection with AIC model selection criterion; for admissions (stage 1)—backward variable selection method and SBC model selection criterion, (stage 2)—random forest.

AUC, area under the receiver-operator characteristic curve; LOS, length of stay; MSE, mean squared error; RPHC, rank probability of high cost; RUB, resource utilisation band; SBC, Schwartz Bayesian Criteria.
office managers looking after patients, and developing and evaluating interventions for patients at different risk levels. 

Simpler models using three easy-to-obtain variables (age, gender and LOS; table 4) had better performance than models that used RUB alone, but were inferior to models that used RUB, age and gender. An important consideration is whether the improved performance of the ACG System over a simpler model justifies the costs needed to implement the ACG System. This decision is a strategic one that depends on not only technical efficiency, but other factors such as cost and how and what the healthcare organisations intend to use their segmentation system for.

Our results suggest that machine learning methods using ACG System model markers confer only marginal improvements in predictive ability, suggesting that it might not be necessary to add such approaches to most models. However, in the case of prediction of the absolute costs, the machine learning approach improves substantially on the R² (from 0.237 to 0.385).

Key advantages of this study include the use of a diverse, randomly selected hospital population, with prospective follow-up of healthcare costs and admissions. Further, by only using data that were available to the hospital, this study reflects real-world use conditions. One of the main limitations of our study is that we looked only at a limited number of specific outcomes (primarily costs and admissions), without considering other outcomes that might be meaningful to patients such as functional ability. Further, we could not account for utilisation-related biases. For example, we could not include patients who died in the community without seeking care at our hospital. Our study does not account for costs that might be accrued by patients in other settings, such as the primary care or private hospital settings. Finally, our study population is limited to adults and excludes children and women receiving pregnancy and gynecology care.

In conclusion, we show that a population stratification and risk prediction tool can be used to appropriately segment populations in a hospital patient population with incomplete healthcare utilisation data. Appropriate implementation of such a tool that may support clinical decision-making and population health management is important.

Contributors JKT—analyses, data interpretation, manuscript drafting and review. XZ—data collection and analysis, and manuscript drafting. DYC—study design, data interpretation and manuscript review. IF—data interpretation, clinical input and manuscript review. CSW—data interpretation, clinical input and manuscript review. JT—study design and data interpretation. SCL—data interpretation and manuscript review. EFS—data interpretation, clinical input and manuscript review. WYL—study design, data analysis, manuscript drafting, review, and is the guarantor of the overall content. All authors have read and approved the final manuscript.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not required.

Ethics approval Ethics approval was obtained from the National Healthcare Group (NHG) Domain Specific Review Board prior to initiating this study (NHG DSRB reference number: 2019/00007). As all participant data were de-identified, a waiver for participant consent was also obtained.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. Data cannot be shared publicly because of the National Health Group Data Protection Policy. Should there be interest in accessing the data to check our analyses, requests can be addressed to the Tan Tock Seng Hospital Office of Clinical Epidemiology, Analytics and Knowledge (OCEAN) at 16 Jalan Tan Tock Seng, NCID Building, Level 4, Singapore 308442. This office serves as our institution’s single point of contact for all data requests.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iD Joshua Kuan Tan http://orcid.org/0000-0002-4518-0477

REFERENCES

3. Hibbard JH, Greene J, Sacks RM, et al. Improving population health management strategies: identifying patients who are more likely to be users of avoidable costly care and those more likely to develop a new chronic disease. Health Serv Res 2017;52:1297–300.


Department of Health Policy and Management at The Johns Hopkins University Bloomberg School of Public Health. The johns hopkins ACG® system version 12.0 user documentation; 2019.


StataCorp. Stata statistical software: release 14. College Station, TX: StataCorp LP, 2015.


