


BMJ Open Predicted quality benefits of achievable performance benchmarks of chronic heart failure care in China: results from a nationwide observational study

Chang Yin,¹ Xi Li,² Chao Wang,² Jingkun Li,² Xiaoqiang Bao,² Qiuju Zhang,² Yupeng Wang,² Xudong Ma ,³ Meina Liu⁴

To cite: Yin C, Li X, Wang C, *et al*. Predicted quality benefits of achievable performance benchmarks of chronic heart failure care in China: results from a nationwide observational study. *BMJ Open* 2020;**10**:e036786. doi:10.1136/bmjopen-2020-036786

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2020-036786>).

CY and XL contributed equally.

Received 03 January 2020
Revised 05 August 2020
Accepted 14 August 2020



© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to

Dr Xudong Ma, Medical Quality Evaluation Office, National Health Commission of the People's Republic of China, Beijing, China; maxudong80@163.com and Professor Meina Liu, Biostatistics, Harbin Medical University, Harbin, China; lumeina369@163.com

ABSTRACT

Objectives This study aimed to set a data-driven achievable performance benchmark, explore the process–outcome association and speculate about the net gain in quality improvement with benchmarking.

Design Observational study.

Setting Patient survey conducted at 466 secondary and tertiary hospitals across 31 provinces, autonomous regions and municipalities in China.

Participants 183 334 patients diagnosed with chronic heart failure (CHF) who were treated at 466 Chinese hospitals from January 2011 through May 2017.

Primary independent variables Hospital process composite performance (HPCP).

Secondary independent variables Patient-level and hospital-level characteristics.

Primary outcome measure Patients getting better or recovered after treatment, in-hospital mortality, length of hospital stay (LOS) and medical cost.

Methods HPCP was calculated using denominator-based weights. Mixed random-intercept models were used to evaluate the contributions of HPCP on patient outcomes and to speculate quality improvement after adjusting HPCP to benchmark level.

Results When all hospitals were to operate at the benchmark level, the proportion of patients getting better or recovered after treatment would increase in most hospitals, particularly those with low baseline rates. However, there was no evidence for lowering in-hospital mortality, significant savings in cost or shortening LOS.

Conclusions Increasing the adherence rate of CHF care and closing the gap in HPCP between hospitals have important implications for improving patient condition.

INTRODUCTION

Chronic heart failure (CHF) remains one of the most common causes of hospitalisation and the most expensive medical condition within Medicare.¹ Despite extensive clinical trial evidence and recommendations in the national guidelines of CHF, there is a significant lag between the incorporation of clinical trial evidence into professional guidelines and delivery of evidence-based care.^{2–5}

Strengths and limitations of this study

- Benchmarking provides performance targets for improvement and promotes analysis and emulation of those providers achieving ‘excellent performance’.
- Using a nationwide chronic heart failure (CHF) medical record data set, we set data-driven achievable performance benchmarks.
- Mixed random-intercept models were used to evaluate the contributions of hospital process composite performance on patient outcomes.
- The speculation of the net gains and trends in quality improvement with benchmarking may provide additional motivation for compliance with clinical practice guidelines.
- Although this was exclusively a data-based estimation, it shed light on the desired effect of ensuring high-quality care for CHF.

The proportion of patients with CHF with favourable outcomes after treatment varies significantly among Chinese medical centres, and whether this variation comes from differences in hospital performance in CHF care has not been well characterised. Findings on the associations between adherence to CHF care and patient outcomes are inconsistent, and the associations may change as the process performance is substantially being improved.^{6–10} More importantly, whether and to what extent the predicted process-adjusted quality improvement in terms of patient outcomes would be achieved when the hospital process performance improved have not been addressed. The answer to this problem may provide additional motivation for promoting the implementation of clinical practice guidelines and patients’ adherence to treatment recommendations.

Benchmarking performance is generally considered to be an important tool for quality improvement.^{11–13} Setting performance

targets can facilitate quality improvement by increasing adherence to evidence-based guidelines and improving patient outcomes. Using a nationwide CHF medical record data set, we aimed to (1) explore the relationship between hospital composite performance of care and patient in-hospital outcomes, and (2) set data-driven achievable performance benchmarks and speculate about the net gains and trends in quality improvement if all hospitals were to operate at the benchmark level.

METHODS

Data source

We obtained data from Medical Care Quality Management and Control System for Specific Disease of China. The system is an ongoing voluntary, continuous, web-based registry and was designed to collect and manage data on quality indicators (QIs). Each QI has an explicit definition. All registered hospitals were secondary or tertiary hospitals from 31 provinces, autonomous regions and municipalities in China. Since 2015, the national medical services, quality and safety report has been published every year, and the data come from the Medical Care Quality Management and Control System for Specific Disease of China.

Each participating hospital appointed a surgeon or nurse responsible for (supervising) the data reporting in a secured web form. Collectors extracted the necessary information from inpatient medical records, drug charts, discharge summary and assessment sheets. Before the online data reporting, collectors underwent data reporting trainings and competency assessments. Only when the assessment was qualified (the inter-rater reliability reached 95% or more) could the data collectors start online data reporting. At point of entry, values that exceed expected ranges or have logical errors will prompt notification. To minimise error caused by artificial reporting, the information on the front page of inpatient medical records was collected automatically. Once collected, de-identified data were entered and aggregated into an analytical database. Annual feedback on data quality problems and treatment quality would be provided to all registered hospitals.

Study population

We combined hospital-level data and patient-level data according to hospital codes and identified a cohort of 196 498 patients with CHF treated at 682 hospitals from 1 January 2011 through 31 May 2017. The exclusion criteria are shown in online supplemental appendix figure 1. First, patients with no outcome record at discharge were excluded. Second, we deleted observations with aberrant values, including outliers or obvious error value. Third, patients admitted in transfer were excluded. Fourth, patients younger than 18 years at diagnosis or patients with length of hospital stay (LOS) more than 120 days or less than 1 day were excluded. Fifth, to avoid inflation in variance owing to few numbers, hospitals with fewer than

20 admissions overall were excluded. Finally, to ensure comparability of the hospital process composite performance (HPCP), hospitals should have eligible patients of all QIs selected in this study. The final analytical population comprised 183 334 patients with CHF from 466 hospitals. Although 216 hospitals and 13164 patients were excluded according to the above-mentioned criteria, there was no significant difference between the included group and the excluded group (see online supplemental appendix tables 1 and 2).

Study variables

Baseline demographics information, comorbidities, laboratory findings, medication history, healthcare utilisation and in-hospital outcomes were recorded. The main predictor variable of quality of care was healthcare utilisation. Patient-level covariates included year and age at diagnosis, primary payer status, pattern of admission, left ventricular ejection fraction (LVEF), New York Heart Association (NYHA) classification and comorbidities. Hospital-level covariates included geographical region, hospital ownership, hospital level, university affiliated, nurse to bed ratio, health technician to bed ratio, and annual outpatient volume.

Primary payer status was classified as socialised medicine, basic medical insurance, other medical insurance types and self-paying. The reimbursement rates of socialised medicine and basic medical insurance in China were 85%–95% and 50%–65%, respectively, and the reimbursement rates of other insurance types were between socialised medicine and basic medical insurance. Not all hospitalisation cares were covered by insurance. The insurance was a factor that doctors considered when recommending appropriate care to patients. LVEF was divided into three categories: CHF with reduced ejection fraction if ejection fraction was $\leq 40\%$, CHF with preserved ejection fraction if ejection fraction was consistently $\geq 50\%$, CHF with mid-range ejection fraction of 41%–49%,¹⁴ and the missing value category 'Unrecorded'.

Selection of QIs and outcomes

QIs were identified with respect to published scientific evidence, consistency with established clinical practice guidelines and their accurate reflection of quality of care.^{15–17} We also considered data completeness and accuracy. Because we aimed to assess the relationship between HPCP and patient in-hospital outcomes, drugs prescribed at discharge were excluded. Finally, 14 priority QIs were selected in the study:

- ▶ Assessment of left ventricular function (LVF) (QI1–QI3): first chest X-ray examination, first echocardiography, first NYHA classification, 6 min walk test and Killip classification (acute myocardial infarction (AMI)) within 24 hours of admission.
- ▶ Drug prescriptions within 24 hours of admission (QI4–QI7): diuretic and potassium, ACE inhibitors (ACEI) or angiotensin receptor blocker (ARB) for patients with LVEF $\leq 40\%$, β -blockers (metoprolol,

propranolol, atenolol), and aldosterone antagonists (aldactone, eplerenone) for patients with NYHA grade III–IV and LVEF <40%, or patients suffering from CHF after AMI with LVEF <40% or Killip grade III–IV.

- ▶ Drug prescription during hospitalisation (QI8–QI11): diuretic and potassium, ACEI or ARB, β -blockers, and aldosterone antagonists during hospitalisation.
- ▶ Health education (QI12–QI13): patient assessment after admission, including cardiac function and self-care assessment, patient symptoms description, smoking and drinking cessation counselling, and psychological counselling; hospitalisation and discharge instructions regarding medications, diet, weight monitoring, worsening of symptoms, follow-up appointment and activity.
- ▶ Risk assessment at discharge (QI14), including chest X-ray examination, echocardiography, NYHA classification, 6 min walk test and Killip classification (AMI), biomarker detection, and risk factor reassessment.

Patients eligible for inclusion in calculations associated with each individual QI included only those who met the criteria for each specific therapy and for whom there were no contraindications, intolerance or other documented reasons to explain why the indicated therapy was not provided. The detailed QI specifications have previously been published.^{18 19} If the physician felt that the patient was not indicated for a certain treatment, he would not recommend it.

The outcomes of patients with CHF were classified by the attending physician. Patient outcomes at discharge were classified into five categories: recovered, getting better, unchanged or worse, died, and unable to determine. ‘Recovery’ was defined as complete restoration of cardiac function and the ability to carry out all usual duties and activities. ‘Getting better’ was defined as the improvement of symptoms and the ability of patients to take care of their own affairs without assistance. ‘Unchanged or worse’ meant that the condition of HF did not change or even worsen after treatment. Patients without a record of outcome due to abandonment of further treatment or other reasons were treated as ‘unable to determine’. The primary outcomes of interest in this study were patients getting better or recovered after treatment. We classified patients who were recovered or getting better into the condition-improved group and the rest into the unimproved group. Other outcome measures included in-hospital mortality, LOS and medical costs for all hospital services and medication.

Statistical analysis

We calculated the adherence to individual QIs by dividing the sum of eligible patients who were provided correct care (numerator) by the total care opportunities (denominator). HPCP was calculated using denominator-based weights, where the weight applied to QI was the ratio of its denominator to the sum of all QIs’ denominators.^{20 21} Patients found to have a clinical care contraindication or those who lacked

information on care eligibility were excluded from both the numerator and the denominator. Greater HPCP meant better adherence to QIs. We divided hospitals into equal quantiles by ranking HPCP scores. Variations in adherence to individual QIs were reported across quality quartiles. Baseline demographics characteristics, comorbidities, in-hospital outcome and hospital characteristics were compared among hospital performance quartiles using χ^2 or Kruskal-Wallis H tests. Fisher’s exact test was used when indicated to compare infrequent events.

Differences between unadjusted and adjusted in-hospital outcomes across hospitals were assessed by mixed random-intercept models with patient-level and hospital-level factors as fixed effects and a random intercept for hospitals. The HPCP score was added as a continuous predictor variable and ORs were reported per 10% increment in HPCP score. We used hierarchical logistic regression to calculate the risk-adjusted condition-improved rate, normal mixed random-intercept models to assess the relationship between HPCP and log-transformed medical cost and LOS, and Poisson general linear mixed model to calculate the risk-adjusted in-hospital mortality.

We used the ‘pared-mean’ approach to set the composite performance benchmark, defined as the average performance of the subset of those hospitals with the highest HPCP.^{11 13} We ranked hospitals in descending order of HPCP. The benchmarking subset included the top-ranked hospitals down to the point whereby at least 10% of the eligible patient pool across all hospitals were selected. Relying on the process–outcome association and using self-control method, we speculated about the net gains and trends in quality improvement of condition-improved rates if all hospitals were to operate at the benchmark level.²² From the full covariate-adjusted hierarchical model, the effect of HPCP was taken as random effect. For the entire study population, we used the hierarchical logistic model to predict their risk-adjusted condition-improved rates given their fixed covariates and the weighted mean random effect of hospitals in the benchmarking subset.

Sensitivity analyses were conducted by dividing hospitals into equal decile groups based on sequentially ordered HPCP scores. We calculated the composite process performance score from the decile level and took the best-performing group (hospitals in the top 10% ranking orders) as the benchmark. All analyses were performed using SAS V.9.3.

Patient and public involvement

Patients or the public were not involved in the design, or conduct, or reporting, or dissemination plans of our research.

RESULTS

Performance on individual indicators and HPCP

The overall proportion of patients eligible for individual CHF indicators ranged from 25.51% (46 767 in 183 334) of ACEI or ARB to 74.97% (137 449 in 183 334) of chest X-ray, except for five full eligible indicators (table 1).

Table 1 Variation in adherence to individual quality indicators by quality quartiles*

Eligible patients (% received)	Overall (N=183 334)	Lowest quality (n=57 099)	Quartile 2 (n=53 697)	Quartile 3 (n=39 596)	Highest quality (n=32 942)
Assessment of left ventricular function					
QI1: chest X-ray	137 449 (40.33)	34 262 (15.48)	40 986 (37.97)	31 426 (46.11)	30 775 (65.24)
QI2: echocardiography	183 334 (54.98)	56 981 (28.41)	53 815 (57.06)	39 596 (65.89)	32 942 (84.44)
QI3: NYHA classification	183 334 (54.56)	56 981 (17.33)	53 815 (58.17)	39 596 (72.85)	32 942 (91.06)
Drug prescription within 24 hours of admission					
QI4: diuretic and potassium	125 942 (58.80)	33 247 (21.90)	38 110 (60.37)	28 560 (70.56)	26 025 (90.73)
QI5: ACEI/ARB	46 767 (85.53)	6 337 (86.70)	15 434 (79.27)	13 566 (88.12)	11 430 (90.28)
QI6: β -blockers	67 734 (75.94)	7 858 (70.35)	23 029 (71.23)	17 642 (78.58)	19 205 (81.44)
QI7: aldosterone antagonists	64 858 (92.82)	5 400 (92.39)	18 169 (90.56)	17 789 (92.20)	23 500 (95.13)
Drug prescription during hospitalisation					
QI8: diuretic and potassium	119 640 (93.46)	14 483 (91.69)	38 907 (90.07)	34 605 (95.58)	31 645 (96.12)
QI9: ACEI/ARB	126 978 (71.25)	20 348 (70.23)	41 483 (66.60)	34 201 (72.56)	30 946 (76.70)
QI10: β -blockers	113 923 (66.36)	13 496 (67.75)	37 498 (63.26)	32 671 (64.87)	30 258 (71.19)
QI11: aldosterone antagonists	116 035 (86.43)	13 722 (84.66)	37 918 (80.67)	33 182 (89.67)	31 213 (90.76)
Health education					
QI12: assessment for patient	183 334 (59.25)	56 981 (20.76)	53 815 (61.90)	39 596 (81.74)	32 942 (94.47)
QI13: hospitalisation and discharge instructions	183 334 (57.78)	56 981 (20.30)	53 815 (59.97)	39 596 (79.28)	32 942 (93.21)
QI14: risk assessment	183 334 (37.10)	56 981 (9.70)	53 815 (35.55)	39 596 (45.41)	32 942 (77.03)

*Hospitals were divided into equal quartiles according to hospital process composite performance. The first quartile was regarded as the lowest quality group and the fourth quartile was regarded as the highest quality group. For detailed definitions of all QIs, refer to the 'Selection of QIs and outcomes' section.

ACEI, ACE inhibitors; ARB, angiotensin receptor blocker; NYHA, New York Heart Association; QI, quality indicators.

Overall adherence rate of the eligible population varied greatly by QIs. Drug prescription of diuretic and potassium during hospitalisation had the highest adherence (93.46%). The lowest adherence was observed in risk assessment at discharge (37.10%). The overall composite performance was 64.98%. The median HPCP was 69.09% (IQR, 50.59%–79.45%). Except for drug prescription, the adherence difference of QIs between the highest quality group and the lowest quality group exceeded 50%. Adherence rate of health education varied significantly among hospital quartiles from 20% in the lowest quality group to more than 90% in the highest quality group. The HPCP of hospitals in the benchmarking subset was greater than 85.27% (maximum value, 97.07%). The benchmarking level of performance was 90.07%, and there was 25.09% room for HPCP improvement with benchmarking.

Patient-level and hospital-level characteristics among quality quartiles

Table 2 lists the distribution of patient-level characteristics among quality quartiles. Of the entire study population, the median age at diagnosis was 73 years (IQR, 64–80

years), and 33.1% were admitted in emergency. Over 90% of the patients benefited from the health insurance system. Of the patients, 17.84% were diagnosed with CHF with reduced ejection fraction, and 80.18% of them were patients with NYHA grade III–IV. Nearly one in three patients had other coexisting conditions and over 21.00% had more than one complication. Compared with hospitals in low-quality quartiles, the high-quality quartiles had a significantly higher rate of coexisting conditions. Hospitals enrolled in this study were basically tertiary and owned by the government (see online supplemental appendix table 3). Non-university-affiliated hospitals were in the majority, accounting for 72.75%. Over half of the hospitals were in eastern China. There was no significant difference in hospital characteristics among quartiles.

Unadjusted and adjusted patient outcomes at discharge

Unadjusted median medical cost and LOS were ¥8511 (IQR, 5617–13692) and 9 (IQR, 7–13) days, respectively. The total proportion of condition-improved patients was 69.08% and varied from 37.92% in the lowest quality group to 93.93% in the highest quality group ($p < 0.0001$

Table 2 Patients' characteristics by quality quartile*

Patients' characteristics	Overall	Lowest quality	Quartile 2	Quartile 3	Highest quality	P value
Patients, n	183 334	57 099	53 697	39 596	32 942	
Age						<0.0001
<60	18.97	19.90	17.53	19.43	19.14	
60–70	23.1	23.10	23.17	22.39	23.80	
70–80	34.76	34.58	35.77	34.01	34.32	
>80	23.18	22.42	23.52	24.16	22.74	
Primary payer status						<0.0001
Socialised medicine	4.18	3.59	2.92	6.34	4.65	
Basic medical insurance	79.43	80.62	78.77	79.75	78.08	
Other insurance types	7.14	6.49	9.74	5.75	5.7	
Self-paying	9.25	9.3	8.57	8.16	11.57	
Patterns of admission						<0.0001
Emergency	33.11	31.05	33.63	34.34	34.35	
Outpatient	66.89	68.95	66.37	65.66	65.65	
Ejection fraction						<0.0001
CHFrEF	17.84	7.53	17.39	21.71	31.78	
CHFPeEF	26.21	15.51	28.89	30.12	35.71	
CHFmrEF	10.51	5.29	10.27	13.43	16.45	
Unrecorded	45.44	71.67	43.45	34.74	16.05	
NYHA classification						<0.0001
Grade I	0.21	0.17	0.24	0.22	0.24	
Grade II	3.08	3.17	3	3.62	2.42	
Grade III	32.4	13.35	39.19	36.9	48.96	
Grade IV	23.55	9.61	20.6	34.54	39.32	
Unrecorded	40.76	73.71	36.98	24.73	9.06	
Year						<0.0001
2011	9.32	9.29	9.42	10.58	7.67	
2012	15.12	15.01	15.54	17.99	11.19	
2013	19.03	18.28	15.24	22.73	22.06	
2014	16.51	14.71	16.84	18.13	17.14	
2015	15.14	17.06	16.26	9.95	16.2	
2016	16.64	18.53	16.39	13.48	17.56	
2017	8.25	7.11	10.31	7.14	8.19	
Coexisting condition						
Diabetes mellitus	9.46	2.33	9.51	10.59	20.37	<0.0001
Myocardial infarction	9.02	2.07	8.7	9.48	21.05	<0.0001
Pneumonia	6.7	1.22	7.37	7.64	13.97	<0.0001
Coronary heart disease	6.66	1.42	4.82	9.5	15.31	<0.0001
Renal failure	4.11	0.85	4.4	4.34	8.98	<0.0001
COPD	5	1.2	5.78	5.24	10.03	<0.0001
Arrhythmia	5.82	1.46	5.93	5.44	13.68	<0.0001
Haematological disease	9.36	2.13	9.62	8.74	22.35	<0.0001
Cerebrovascular disease	4.04	0.88	4.33	3.82	9.33	<0.0001
Rheumatic valvular heart disease	3.55	0.7	3.36	4.21	8	<0.0001
Congestive heart failure	18.55	4.02	18.28	21.17	41	<0.0001

Continued

Table 2 Continued

Patients' characteristics	Overall	Lowest quality	Quartile 2	Quartile 3	Highest quality	P value
In-hospital mortality	1.05	0.82	1.29	1.23	0.84	<0.0001
Treatment outcome						<0.0001
Unimproved	30.92	62.08	25.91	13.47	6.07	
Improved	69.08	37.92	74.09	86.53	93.93	
Medical costs, median (¥)	8511	8633	8548	8392	8411	<0.0001
LOS, median, day	9	9	9	10	9	<0.0001

*Patients are divided into quartiles based on hospital ranks sorted by hospital process composite performance scores. The first quartile was regarded as the lowest quality group and the fourth quartile was regarded as the highest quality group. All percentages are based on non-missing values.

CHFmrEF, congestive heart failure with mid-range ejection fraction; CHFpEF, congestive heart failure with preserved ejection fraction; CHFrEF, congestive heart failure with reduced ejection fraction; COPD, chronic obstructive pulmonary disease; LOS, length of stay; NYHA, New York Heart Association.

for trend; figure 1). Of the patients, 1.05% died in hospital. After adjusting patient-level and hospital-level characteristics, every 10% increment in HPCP score was associated with an overall 43% increase in the proportion of condition-improved patients (OR 1.43, 95% CI 1.36 to 1.50). There was little evidence of an association between in-hospital mortality and HPCP (OR 1.06, 95% CI 0.99 to 1.13). Medical costs and LOS were negatively correlated with HPCP, but not significantly ($p>0.05$; table 3).

Predicted quality improvement after adjusting HPCP to benchmark level

There was a significant strong correlation between hospital composite performance on CHF care and

condition-improved rates (Spearman correlation coefficient=0.8086, $p<0.0001$). When all hospitals were to operate at the benchmark level, risk-adjusted condition-improved rate increased in most hospitals and hospitals with lower baseline rate had a larger increase (figure 2A). The overall risk-adjusted condition-improved rate was 82.91% with a net gain of 13.83%. Out of 466 hospitals, 453 demonstrated significant change in risk-adjusted condition-improved rate after adjusting HPCP to benchmark level ($p<0.05$; figure 2B). Of the hospitals, 32.75% (146 in 466) showed a downward trend and 65.89% (307 in 466) showed an upward trend. Hospitals were divided into equal decile groups in the sensitivity analysis, but

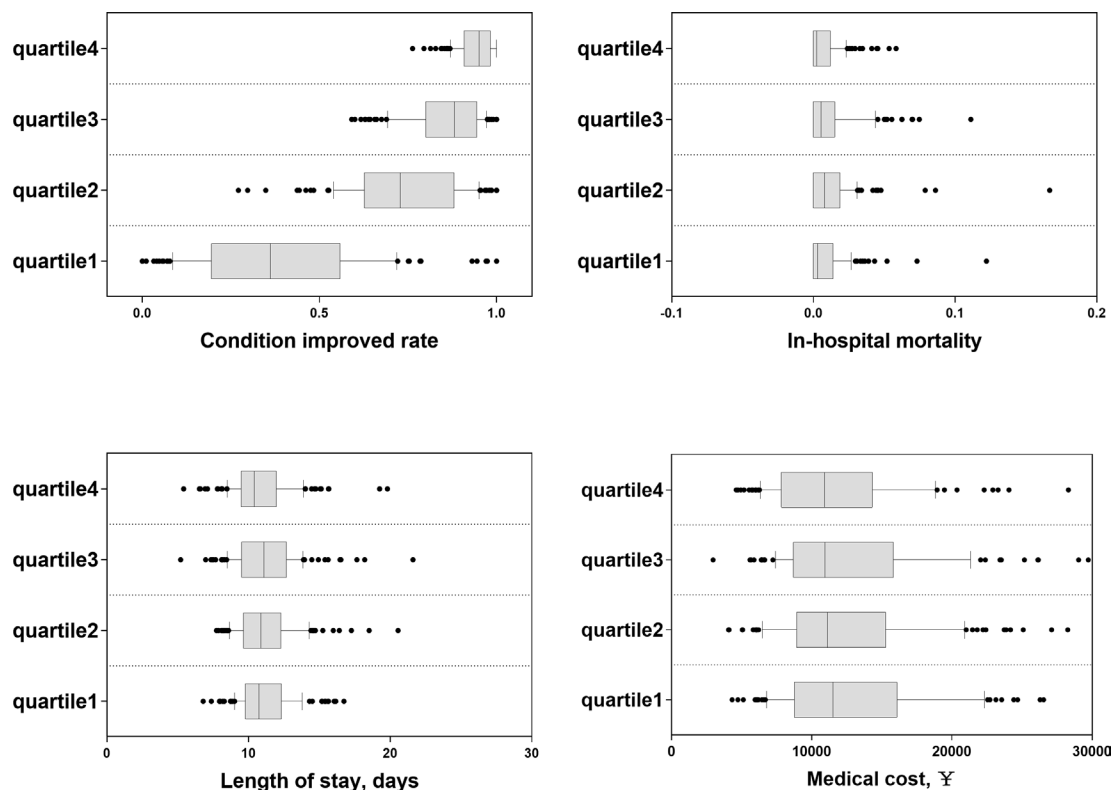


Figure 1 Distributions of in-hospital outcomes among quality quartiles.

Table 3 Unadjusted and adjusted patient outcomes at discharge

Patient outcomes	Risk-adjusted rate or adjusted median				Regression coefficient (95% CI)†	
	Lowest quality	Quartile 2	Quartile 3	Highest quality	Unadjusted	Adjusted
Condition-improved	61.48	67.81	67.04	69.37	0.648 (0.605 to 0.692)*	0.356 (0.306 to 0.406)*
In-hospital mortality	0.75	0.59	0.69	1.03	1.052 (0.994 to 1.114)	0.060 (−0.003 to 0.123)
Medical cost	9241	9029	8396	8332	−0.007 (−0.023 to 0.008)	−0.004 (−0.016 to 0.008)
Length of stay	10.45	10.62	11.15	10.43	−0.035 (−0.127 to 0.058)	−0.062 (−0.152 to 0.027)

*P<0.05.

†Regression coefficient (95% CI) of hospital process composite performance scores. Adjusted for age group, primary payer status, left ventricular ejection fraction, New York Heart Association classification, pattern of admission, year, comorbidities, nurse to bed ratio, health technician to bed ratio, geographical region, ownership, hospital level, university affiliated, and annual outpatient volume. The first quartile was regarded as the lowest quality group and the fourth quartile was regarded as the highest quality group.

the result of benchmarking analysis at the decile level was similar to that at the hospital level. Most of the decile groups had an increased risk-adjusted condition-improved rate through the benchmarking process and the rising degree ranged from 0.04% in the highest decile to 46.15% in the lowest decile (see online supplemental appendix figure 2).

DISCUSSION

In recent years, there has been an increasing interest in performance measurement and quality improvement in healthcare due to the identification of gaps in care and the rising healthcare expenditures.¹² Establishing attainable benchmarks for performance has been proven to drive quality improvement by providing a framework for continuous improvement.²³ In the current study, we explored the associations between adherence to CHF care and patient in-hospital outcomes, and we set a data-driven achievable performance benchmark and speculated about the net gain in quality improvement with benchmarking. Significantly positive association was observed between performance on guideline-recommended care and improvement of patient condition after treatment. However, there was little evidence that better process

performance could reduce in-hospital mortality, significantly save costs or shorten hospital stay. When the composite performance of all hospitals was improved to the benchmark level, we observed a significant improvement in the proportion of patients getting better or recovered.

The important goal of the benchmark-setting was to aid in the spread of ‘excellent performance’ achieved by a few superior providers until they became ‘average performance’ by the majority. Data-driven benchmarking set objective targets by ranking hospital performance on care and identifying the best-performing hospitals that represented the top decile of the eligible population. Compared with benchmarks selection of the mean performance or the top 10th percentile of providers or institutions, focusing on the top 10% of the total eligible population enabled all high-performing providers to contribute to the target and ensured that high-performing providers with low case volumes did not unduly influence the benchmark levels, thereby ensuring that benchmarks were both realistic and attainable.^{11 13}

Overall risk-adjusted condition-improved rate of patients with CHF increased more than 10% when we presumed that the composite performance of all hospitals was

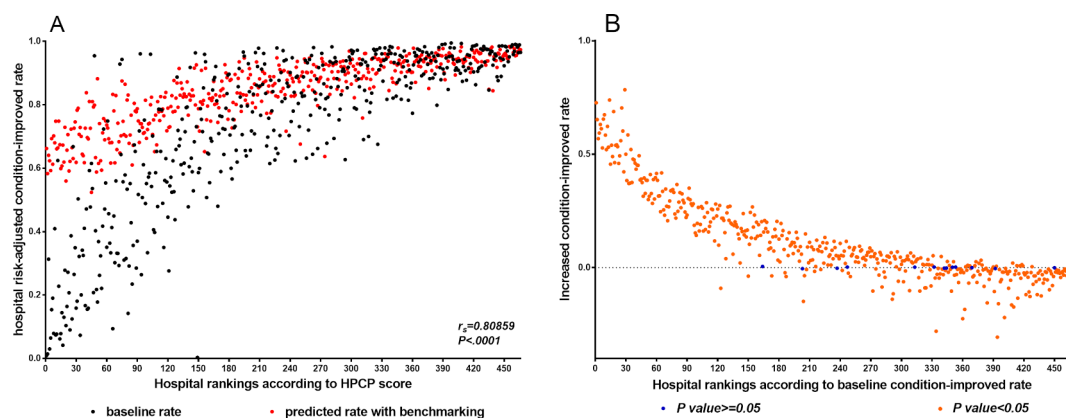


Figure 2 Risk-adjusted condition-improved rate (A) and predicted net gain in quality improvement (B) after adjusting HPCP to the benchmark level. The models have been adjusted for age group, primary payer status, left ventricular ejection fraction, New York Heart Association classification, pattern of admission, year, comorbidities, nurse to bed ratio, health technician to bed ratio, geographical region, ownership, hospital level, university affiliated, and annual outpatient volume. HPCP, hospital process composite performance.



improved to benchmark level. Moreover, lower baseline rate had a larger increase. Large room for improvement in both HPCP and the proportion of condition-improved patients suggested that many patients did not receive optimal care, and this in turn was responsible for many avoidable poor prognosis or deaths. The speculation of the net gains and trends in quality improvement with benchmarking may provide additional motivation for compliance with clinical practice guidelines. Increasing the adherence rate of CHF care and closing the gap in HPCP between hospitals had important implications for improving patient condition. Meanwhile, we observed decline in some hospitals when given their fixed covariates and the hospital effect of the benchmarking subset. The regression contribution (linear predictor) of HPCP accounted for one-third of the hospital effect, indicating that there were still unmeasured/immeasurable factors in case-mix that might have affected the outcome irrespective of hospital performance on treatment.

In our study, the quality difference across the quartiles was mainly from QIs related to LVF assessment and health education. HF-specific health education provides an inexpensive means that may improve physician adherence to treatment goals, patient compliance with instructions as well as management of comorbid conditions.²⁴ However, many hospitals in China have not paid enough attention to risk assessment and health education of patients with CHF, and providers have underestimated its impact on patient outcomes. The proportion of coexisting diseases in high-quality quartiles was higher than that in low-quality quartiles, possibly because patients who were more ill were generally sent to hospitals with better medical equipment and staffing. Besides, physicians paid more attention to risk assessment and health education in patients with more severe conditions. The adherence rate of most QIs selected in this study still has a large room to improve. Special emphasis should be given to QIs related to LVF assessment and health education in quality improvement of CHF care. Also, medicines management at home is important for quality improvement of CHF care. Half to two-thirds of all rehospitalisations have been associated with poor adherence to medication.^{25 26}

Our finding showed that good performance on CHF care could best motivate positive change in conditional improvement, rather than other patient outcomes in the study. The process–outcome association has been analysed for other medical conditions, such as cardiac arrest, renal cell carcinoma, AMI, acute coronary syndrome and so on.^{22 27–29} In most of the literature, the outcome measure was risk-adjusted mortality. Findings on the associations between hospital process performance and clinical outcomes were inconsistent.^{6 7 30} Most researchers found that hospital process performance had little relationship to risk-adjusted mortality rates.^{9 10 31} Even if a significant correlation was observed between risk-adjusted mortality and certain process measures, hospital process performance could only explain a small part of the variation in mortality rates.^{6 32} Our finding was in agreement with

previous studies that showed good process performance could not reduce in-hospital mortality. In our study, in-hospital mortality of patients with CHF was low with a declining trend, and many interventions in clinics were palliative rather than curative.² The most direct clinical manifestation of the therapeutic effect is alleviation of clinical symptoms and improvement of patient condition after treatment. For chronic diseases, we should focus on function improvement and life quality after treatment.

Our study is a large nationally representative retrospective case record review conducted in China. The nationwide data set used in our study helped to ensure the reliability and validity of the association between adherence and outcome. Although this was exclusively a data-based estimation, it shed light on the desired effect of ensuring high-quality care for CHF. Our study has several limitations. First, the data were limited to patient-level and hospital-level information. Information on individual physicians requires a more efficient and comprehensive data collection and reporting system. Second, the data were abstracted from medical record reviews, and imperfect documentation may have diluted the observed associations. Third, because of the absence of follow-up, we could not figure out the contribution of better adherence to long-term outcome.

CONCLUSION

Hospital composite performance of CHF care was significantly associated with the proportion of patients getting better or recovered after treatment. Examining the process–outcome association provides evidence for making management decisions on CHF care and establishing actionable targets for quality improvement. Increasing the compliance rate of CHF care and closing the gap in HPCP have important implications for quality improvement efforts.

Author affiliations

¹Information Center, National Institute of Hospital Administration, Beijing, China

²BioStatistics, Harbin Medical University, Harbin, Heilongjiang, China

³Medical Quality Evaluation Office, National Health Commission of the People's Republic of China, Beijing, China

⁴Public Health College, Harbin Medical University, Harbin, China

Acknowledgements The authors thank the reviewers and editors of *BMJ Open* for providing helpful comments to improve the manuscript.

Contributors CY and XL were the principal investigators and contributed to study design, data analyses and interpretation of findings, and wrote the manuscript. CW and JL contributed to study design, data analyses and interpretation of findings. XB, QZ and YW contributed to study design, interpretation of findings and article revision. XM and ML contributed to interpretation of findings, article revision and study supervision. All authors read and approved the final version of the manuscript.

Funding This work was supported by the National Natural Science Foundation of China (81573255 to ML).

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data may be obtained from a third party and are not publicly available. The data that support the findings of this study are available from

the Medical Care Quality Management and Control System for Specific Disease of China, but restrictions apply to the availability of these data, which were used under licence for the current study and so are not publicly available.

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

Xudong Ma <http://orcid.org/0000-0002-9924-0682>

REFERENCES

- Writing Group Members, Mozaffarian D, Benjamin EJ, *et al*. Heart disease and stroke Statistics-2016 update: a report from the American heart association. *Circulation* 2016;133:e38.
- Stout KK, Broberg CS, Book WM, *et al*. Chronic heart failure in congenital heart disease: a scientific statement from the American heart association. *Circulation* 2016;133:770.
- Hernandez AF, Hammill BG, Peterson ED, *et al*. Relationships between emerging measures of heart failure processes of care and clinical outcomes. *Am Heart J* 2010;159:406–13.
- Fonarow GC, Albert NM, Curtis AB, *et al*. Associations between outpatient heart failure process-of-care measures and mortality. *Circulation* 2011;123:1601–10.
- Fonarow GC, Albert NM, Curtis AB, *et al*. Improving evidence-based care for heart failure in outpatient cardiology practices: primary results of the registry to improve the use of evidence-based heart failure therapies in the outpatient setting (improve HF). *Circulation* 2010;122:585–96.
- Pitches DW, Mohammed MA, Lilford RJ. What is the empirical evidence that hospitals with higher-risk adjusted mortality rates provide poorer quality care? A systematic review of the literature. *BMC Health Serv Res* 2007;7:91.
- Maeda JLK. Evidence-Based heart failure performance measures and clinical outcomes: a systematic review. *J Card Fail* 2010;16:411–8.
- Dy SM, Chan KS, Chang H-Y, *et al*. Patient perspectives of care and process and outcome quality measures for heart failure admissions in US hospitals: how are they related in the era of public reporting? *Int J Qual Health Care* 2016;28:522–8.
- Werner RM, Bradlow ET. Relationship between Medicare's hospital compare performance measures and mortality rates. *JAMA* 2006;296:2694.
- Patterson ME, Hernandez AF, Hammill BG, *et al*. Process of care performance measures and long-term outcomes in patients hospitalized with heart failure. *Med Care* 2010;48:210–6.
- Weissman NW, Allison JJ, Kiefe CI, *et al*. Achievable benchmarks of care: the ABCs of benchmarking. *J Eval Clin Pract* 1999;5:269–81.
- Ettorchi-Tardy A, Levif M, Michel P. Benchmarking: a method for continuous quality improvement in health. *Healthc Policy* 2012;7:E101–19.
- Kiefe CI, Weissman NW, Allison JJ, *et al*. Identifying achievable benchmarks of care: concepts and methodology. *Int J Qual Health Care* 1998;10:443–7.
- Lam CSP, Solomon SD. Fussing over the middle child: heart failure with Mid-Range ejection fraction. *Circulation* 2017;135:1279–80.
- Chinese Society of Cardiovascular Medicine. Guidelines for diagnosis and treatment of heart failure in China 2014 [in chinese]. *Chinese Journal of Cardiology* 2014;42:3–10.
- Yancy CW, Jessup M, Bozkurt B, *et al*. 2013 ACCF/AHA Guideline for the Management of Heart Failure: Executive Summary : A Report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. *J Am Coll Cardiol* 2013;62:1495–539.
- Measuring and improving quality of care : A report from the American heart Association/American college of cardiology first scientific forum on assessment of healthcare quality in cardiovascular disease and stroke. *Stroke* 2000;31:1002–12.
- Fonarow GC, Yancy CW, Albert NM, *et al*. Improving the use of evidence-based heart failure therapies in the outpatient setting: the improve HF performance improvement registry. *Am Heart J* 2007;154:12–38.
- Bonow RO, Bennett S, Casey DE, *et al*. ACC/AHA clinical performance measures for adults with chronic heart failure: a report of the American College of Cardiology/American heart association Task force on performance measures (writing Committee to develop heart failure clinical performance measures): endorsed by the heart failure Society of America. *Circulation* 2005;112:1853–87.
- Shwartz M, Ren J, Peköz EA, *et al*. Estimating a composite measure of hospital quality from the hospital compare database: differences when using a Bayesian hierarchical latent variable model versus denominator-based weights. *Med Care* 2008;46:778–85.
- Peterson ED, DeLong ER, Masoudi FA, *et al*. ACCF/AHA 2010 position statement on composite measures for healthcare performance assessment: a report of the American College of Cardiology Foundation/American heart association Task force on performance measures (writing Committee to develop a position statement on composite measures). *Circulation* 2010;121:1780–91.
- Anderson ML, Nichol G, Dai D, *et al*. Association between hospital process composite performance and patient outcomes after in-hospital cardiac arrest care. *JAMA Cardiol* 2016;1:37.
- Thonon F, Watson J, Saghachian M. Benchmarking facilities providing care: an international overview of initiatives. *SAGE Open Med* 2015;3:205031211560169.
- Farmakis D, Filippatos G, Parissis J, *et al*. Structured discharge instructions for hospitalized heart failure patients to improve guideline implementation and patient outcomes. *Int J Cardiol* 2016;220:143–5.
- Ewen S, Baumgarten T, Rettig-Ewen V, *et al*. Analyses of drugs stored at home by elderly patients with chronic heart failure. *Clin Res Cardiol* 2015;104:320–7.
- Ruppar TM, Cooper PS, Mehr DR, *et al*. Medication adherence interventions improve heart failure mortality and readmission rates: systematic review and Meta-Analysis of controlled trials. *J Am Heart Assoc* 2016;5:e002606.
- Lawson KA, Saarela O, Abouassaly R, *et al*. The impact of quality variations on patients undergoing surgery for renal cell carcinoma: a national cancer database study. *Eur Urol* 2017;72:379–86.
- Glickman SW, Ou F-S, DeLong ER, *et al*. Pay for performance, quality of care, and outcomes in acute myocardial infarction. *JAMA* 2007;297:2373–80.
- Peterson ED, Roe MT, Mulgund J, *et al*. Association between hospital process performance and outcomes among patients with acute coronary syndromes. *JAMA* 2006;295:1912–20.
- Mehta RH, Peterson ED, Califf RM. Performance measures have a major effect on cardiovascular outcomes: a review. *Am J Med* 2007;120:398–402.
- Fonarow GC, Abraham WT, Albert NM, *et al*. Association between performance measures and clinical outcomes for patients hospitalized with heart failure. *JAMA* 2007;297:61–70.
- Lingsma HF, Dippel DWJ, Hoeks SE, *et al*. Variation between hospitals in patient outcome after stroke is only partly explained by differences in quality of care: results from the Netherlands stroke survey. *J Neurol Neurosurg Psychiatry* 2008;79:888–94.