


# Population Segmentation Based on Healthcare Needs: Validation of a Brief Clinician-Administered Tool



Jia Loon Chong, MD, PhD<sup>1</sup>, David Bruce Matchar, MD<sup>1,2,3</sup> , Yuyang Tan, MBBS<sup>3</sup>, Shalini Sri Kumaran, MBBS<sup>3</sup>, Mihir Gandhi, PhD<sup>4,5,6</sup>, Marcus Eng Hock Ong, MBBS<sup>1,7</sup>, and Kok Seng Wong, MBBS<sup>3</sup>

<sup>1</sup>Program in Health Services and Systems Research, Duke-NUS Medical School Singapore, Singapore; <sup>2</sup>Department of Medicine (General Internal Medicine), Duke University Medical Center/Durham, NC, USA; <sup>3</sup>Department of Internal Medicine, Singapore General Hospital/Singapore, Singapore; <sup>4</sup>Centre for Quantitative Medicine, Duke-NUS Medical School Singapore, Singapore; <sup>5</sup>Department of Biostatistics, Singapore Clinical Research Institute/Singapore, Singapore; <sup>6</sup>Tampere Center for Child Health Research, Tampere University/Tampere, Finland; <sup>7</sup>Department of Emergency Medicine, Singapore General Hospital/Singapore, Singapore.

**BACKGROUND:** As populations age with increasingly complex chronic conditions, segmenting populations into clinically meaningful categories of healthcare and related service needs can provide healthcare planners with crucial information to optimally meet needs. However, while conventional approaches typically involve electronic medical records (EMRs), such records do not always capture information reliably or accurately.

**OBJECTIVE:** We describe the inter-rater reliability and predictive validity of a clinician-administered tool, the Simple Segmentation Tool (SST) for categorizing older individuals into one of six Global Impression (GI) segments and eight complicating factors (CFs) indicative of healthcare and related social needs.

**DESIGN:** Observational study ([ClinicalTrials.gov](https://clinicaltrials.gov/ct2/show/study/NCT02663037), number NCT02663037).

**PARTICIPANTS:** Patients aged 55 years and above.

**MAIN MEASURES:** Emergency department (ED) subjects (between May and June 2016) had baseline SST assessment by two physicians and a nurse concurrently seeing the same individual. General medical (GM) ward subjects (February 2017) had a SST assessment by their principal physician. Adverse events (ED visits, hospitalizations, and mortality over 90 days from baseline) were determined by a blinded reviewer. Inter-rater reliability was measured using Cohen's kappa. Predictive validity was

evaluated using Cox hazard ratios based on time to first adverse event.

**KEY RESULTS:** Cohen's kappa between physician-physician, service physician-nurse, and physician-nurse pairs for GI were 0.60, 0.71, and 0.68, respectively. Cox analyses demonstrated significant predictive validity of GI and CFs for adverse outcomes.

**CONCLUSIONS:** With modest training, clinicians can complete a brief instrument to segment their patient into clinically meaningful categories of healthcare and related service needs. This approach can complement and overcome current limitations of EMR-based instruments, particularly with respect to whole-patient care.

**TRIAL REGISTRATION:** ClinicalTrials.gov Identifier: NCT02663037

**KEY WORDS:** aging; health services research; psychometrics.

## Abbreviations

CF	complicating factors
CIRB	Centralised Institutional Review Board
ED	emergency department
EMR	electronic medical record
GI	Global Impression
GM	general medical
SGH	Singapore General Hospital
SST	Simple Segmentation Tool

**Prior Presentations** Chong JL, (September, 2018). Matching health and social services to needs in an ageing population: introduction of the Simple Segmentation Tool. Oral presentation at: Singhealth Duke-NUS Scientific Congress, Singapore.

Chong JL, Mihir G, Ong MEH, Matchar DB. (June, 2017). Can clinicians segment patient populations based on global healthcare needs? Inter-rater reliability, convergent validity and predictive accuracy. Poster presented at: Academy Health Conference, New Orleans.

Chong JL, Tan YY, Sri Kumaran S, Wong KS, Matchar DB. (June, 2018). Evaluation of a clinician-administered healthcare needs-based population segmentation tool in a Singaporean tertiary hospital. Poster presented at: Society of General Internal Medicine Meeting, Denver.

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s11606-020-05962-4>) contains supplementary material, which is available to authorized users.

Received April 28, 2020

Revised April 28, 2020

Accepted June 4, 2020

Published online June 30, 2020

J Gen Intern Med 36(1):9–16

DOI: 10.1007/s11606-020-05962-4

© Society of General Internal Medicine 2020

## INTRODUCTION

Worldwide, the proportion of older individuals is increasing due to rising life expectancy and decreasing fertility rates.<sup>1</sup> One consequence of this is an increasing prevalence of chronic diseases. This epidemiological shift has led to a growing recognition that we are not fully meeting needs in an increasingly complex population and has reinforced the notion that health systems must be systematic in meeting key needs to reduce avoidable adverse outcomes.<sup>2–7</sup> However, a disease-

based approach to healthcare resource planning has a tendency to lead to fragmentation of care.<sup>8</sup>

In this context, population segmentation into categories of related service needs is an important foundation for meeting needs in ways that are both effective and sustainable.<sup>7, 9, 10</sup> A planning strategy based on a population segmentation perspective would, for example, facilitate the development of care packages based on sets of similar needs within one segment but differ between segments, resulting in a more integrated approach to addressing diverse healthcare needs<sup>11</sup> and reducing the risk of over- or under-planning of services.<sup>7, 9, 10</sup> For example, the Valchronic integrated care program was found to have led to a reduction in emergency care service use and an increase in patient satisfaction.<sup>9</sup> It may also allow streamlining of programs, such as those involving large multidisciplinary teams, which can be very expensive, especially when they are hospital-based.<sup>12, 13</sup>

Approaches that have been proposed for population segmentation are often proprietary in nature, complex, and focus on risk predictions for healthcare utilization and adverse outcomes.<sup>14</sup> Further, most approaches rely on the electronic medical record (EMR) where key factors known to predict utilization and healthcare needs, such as social support, cannot be obtained reliably.<sup>15, 16</sup>

The “Bridges to Health” segmentation scheme developed by Lynn et al. is an exemplar of a basic needs-based framework<sup>10</sup> aimed at categorizing individuals not by their diseases (e.g., heart failure vs. chronic lung disease) but by the nature of chronic services that would best serve their needs. Specifically, “Bridges to Health” focuses on whether the person has a dominant condition and whether that condition is asymptomatic, symptomatic but stable, or advanced with frequent exacerbations. While the framework has strong clinical face-validity, there are no validated tools to make it operationally useful.

Based on the concepts represented in the “Bridges to Health” framework, we developed the “Simple Segmentation Tool” (SST), a brief instrument used to categorize elderly patients in a clinical setting for purposes of policy planning and evaluation. In order to gain acceptance, the SST needs to be reliable, valid, and easy to use. Thus, the primary aim of our study is (1) to evaluate the inter-rater reliability of the elements of the SST, both between physician pairs and physician-nurse pairs and (2) to assess the validity of the tool in predicting the time to adverse medical outcomes, namely, emergency department (ED) visits, non-elective hospitalizations, and mortality 90 days post-assessment.

## METHODS

### The SST Instrument

The SST<sup>16</sup> (Appendix) enables patient population segmentation based on a clinician rater’s Global Impression (GI) of a

patient’s medical condition relative to a prototypical set of categories.<sup>7, 16</sup> The six GI population segments were adapted from the “Bridges to Health” population segmentation scheme,<sup>10</sup> with a focus on older populations (e.g., age 55 years old and above) as well as long-term health states (i.e., requiring chronic services) (Table A1 of Appendix).

In addition to the core GI, the SST includes eight complicating factors (CFs) (Table A2 of Appendix). These factors were chosen based on findings from a focus group of four experienced clinicians representing general internal medicine, geriatrics, and family medicine to identify factors that could complicate medical management but could be managed through non-medical services; and where published evidence was available to support that factor as a predictor of adverse medical outcomes.<sup>17–23</sup> As shown in Figure A1, CFs were rated on three levels of severity, namely, “low,” “moderate,” or “high”; a “low” level corresponded with the absence of a CF altogether, while “moderate” and “high” levels corresponded with varying degrees of intensity deemed indicative of needs for different types of services. The guidelines for triaging patients into GI and CF levels were placed on the reverse side of the SST instrument (see Fig. A2).

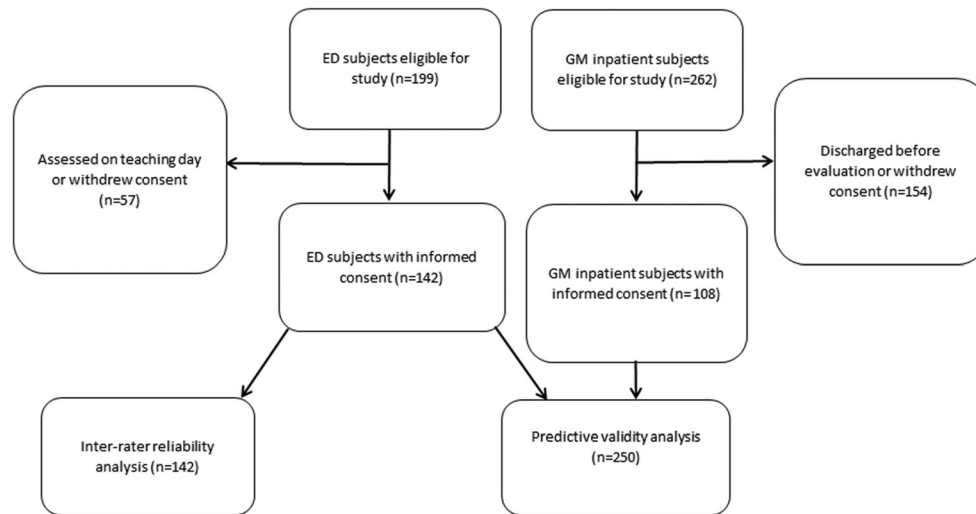
### Subjects

The Singapore General Hospital (SGH) is the largest tertiary acute care hospital in Singapore with over 81,000 hospital admissions and 128,000 ED attendances per year.<sup>24</sup> Subjects were recruited from the ED of SGH between May and June 2016 and the general medical (GM) ward of the Department of Internal Medicine on February 10, 2017 (Fig. 1). ED subjects were eligible if they were Singaporean citizens or permanent residents aged 55 years and above and had not been triaged to the most severe triage urgency category. Subjects were recruited sequentially and informed consent could be obtained either from the subject or their legal representatives. Our sample size for the ED recruitment was 123 based on achieving a predetermined level of inter-rater reliability as measured by Cohen’s kappa.

GM subjects were eligible based on the same criteria except that there was no restriction on severity of their condition as long as they or their legal representatives could provide consent. All patients on the GM ward were assessed on February 10, 2017. The GM study was designed to be an observational study on all patients; based on historical admission rates, at least 100 patients were expected to consent.

### SST Raters

SST raters in the ED consisted of four physicians assigned to provide patient care (service physicians), one independent physician to observe the encounter (observing physician), and two nurses to observe the encounter (see “Procedures” section). The ED raters familiarized themselves with the SST through



**Figure 1** Flowchart of subjects recruited and available for analysis from the emergency department and General Medical Ward.

completion of an online SST tutorial as well as a teaching day during which they had the opportunity to discuss their experience with the SST form and study procedures. Ratings from the teaching days were not included in the analysis. For inter-rater reliability, all ratings were used. For predictive validity, the rating from the service physician was used.

SST raters in the GM ward included 34 physicians. Their training consisted of a SST tutorial where they had the opportunity to clarify SST rating procedures with the study team. In both the ED and GM setting, the provided training modules can typically be completed in less than 15 minutes. For predictive validity, the SST was assessed whenever possible by the physician most directly responsible for the subject's care.

## Procedures

**Ethics Approval.** Both the ED and GM components of this study were approved by the Singapore SingHealth Centralised Institutional Review Board (CIRB) (CIRB/2016/2005 and CIRB/2016/2629 respectively). Written informed consent to participate and to allow access to the patients' EMRs was obtained from participants or their legal representatives.

## Inter-rater Reliability

A research nurse recruited eligible ED patients in the waiting area. Once the subject was brought to a clinic room to be seen by the service physician, a nurse and an observing physician entered the room and silently observed the interaction between the patient and the service physician. The observing physician and nurse were not involved with management of the patient. Consistent with standard clinical practice, all three raters had access to the subject's EMR prior to the index ED visit. After the consultation, each rater completed the SST independently without discussion, resulting in three concurrent SST evaluations per patient.

## Predictive Validity

Predictive validity was assessed by comparing the SST rating for both ED and GM subjects with our outcomes of interest over 90 days after the index SST assessment. The following outcome data were obtained through EMR review: ED visits, non-elective hospital admissions, and mortality within the aforementioned 90-day period. The ED encounter or hospital admission during which the SST was assessed was not counted as an event. The EMR review was performed by an investigator (CJL) without knowledge of the ED or GM encounter itself nor the responses on the SST.

## Statistical Methods

The Cohen's kappa statistics was used to evaluate the degree of inter-rater agreement beyond chance through the concomitant assessments performed in the ED setting. The strength of agreement based on kappa statistics was rated as follows: slight, < 0.20; fair, 0.21–0.40; moderate, 0.41–0.6; substantial, 0.61–0.80; and almost perfect, 0.81–1.00.<sup>25</sup>

Predictive validity of the GI segments and CFs was assessed using Cox proportional hazard models. Hazard ratios for adverse events were computed between the highest and lowest levels of both GI and CF variables, adjusted for age and gender, and censoring for mortality events. For example, hazard rates were compared between the GI categories of "healthy" vs. "short period of decline before dying" (see SST rating page in Fig. A1). Hazard rates for CFs were computed using the differences between "high" and "low" levels of each of the eight variables measured. Kaplan-Meier plots were constructed for the three outcomes stratified by GI category.

## Sample Size Calculation

The sample size for this study was estimated based on the primary objective of assessing the inter-rater reliability of GI

ratings between physician-physician pairs. A sample size of 123 subjects was needed to achieve 80% power (significance level of 0.05) to detect a true kappa value of 0.60 (assuming null value  $k_0 = 0.4$ ) in a test of  $H_0: \text{kappa} = \kappa_0$  vs.  $H_1: \text{kappa} \neq \kappa_0$  when there were six categories with a population distribution equal to 20% (healthy), 50% (chronic conditions, asymptomatic), 25% (chronic conditions, symptomatic), 3% (long course of decline), 1% (limited reserve and serious exacerbations), and 1% (short course of decline before dying).

## Role of the Funding Source

The funding source was not involved in the design of the study, analysis and interpretation of the data, or decision to approve publication of the finished manuscript.

## RESULTS

A total of 199 subjects from the ED were eligible to participate. After excluding subjects who did not consent or were recruited during teaching days, 142 ED subjects were included in analysis. From the GM ward, 262 individuals were approached to participate out of which 108 both consented and received an SST assessment. Thus, 142 individuals were available for assessing inter-rater reliability and 250 subjects were available for assessing predictive validity (142 from the ED and 108 from the GM ward) (Fig. 1). From the ED, twelve subjects with missing data elements (Global Impression and functional assessment) were excluded from analyses requiring those elements.

## Demographics

Of the 142 patients surveyed in the ED, the majority of subjects were between the ages of 55 and 64 (44%) while on the GM ward, the majority were aged 75 or older (57%). In both groups, subjects were equally distributed between men and women and the majority were of Chinese ethnicity (83% in the ED and 69% in the GM ward) (Table 1).

## Distribution by GI and CF Assessments

The ED and GM ward subjects complemented each other in terms of GI segment with the majority in the ED in GI category I or II (76%) and the majority in the GM ward in GI categories II, IV, V, or VI (90%).

While the majority of ED subjects were deemed to have few CFs (e.g., 90% were free of any functional deficits, 99% without disruptive behavioral issues, and 96% without skilled nursing task needs), the GM ward subjects had more CFs implying substantial needs (e.g., 52% with functional deficits that would require external assistance, 71% without coordination of a complex mix of medical services (significant health conditions without a main service provider or multiple non-coordinated providers), and nearly half (44%) with a skilled nursing-type task (e.g., wound or catheter care)).

**Table 1 Patient Characteristics at Baseline**

Variables	Number of ED subjects, N (%)	Number of GM subjects, N (%)
Age		
55–64	62 (43.7)	19 (17.6)
65–74	49 (34.5)	28 (25.9)
>75	31 (21.8)	61 (56.5)
Gender		
Male	73 (51.4)	51 (47.2)
Female	69 (48.6)	57 (52.8)
Race, n (%)		
Chinese	118 (83.1)	75 (69.4)
Indian	8 (5.6)	15 (13.9)
Malay	16 (11.3)	18 (16.7)
Global Impression		
Healthy	20 (14.3)	1 (0.9)
Chronic condition, asymptomatic	86 (61.4)	10 (9.3)
Chronic condition, moderately/seriously symptomatic	29 (20.7)	36 (33.3)
Long course of decline	1 (0.7)	43 (39.8)
Limited reserve and serious exacerbations	2 (1.4)	7 (6.5)
Short period of decline before dying	2 (1.4)	11 (10.2)
Total	140 (100.0)	108 (100.0)
Functional deficit		
None	127 (90.1)	52 (48.2)
Moderate	2 (1.4)	24 (22.2)
High	12 (8.5)	32 (29.6)
Social support		
High	109 (76.8)	56 (51.9)
Moderate	23 (16.2)	32 (29.6)
Low	10 (7)	20 (18.5)
Hospital admissions past 6 months		
None	119 (83.8)	48 (44.4)
1 to 2	21 (14.8)	38 (35.2)
3 or more	2 (1.4)	22 (20.4)
Disruptive behavioral issues		
None	141 (99.3)	93 (86.1)
Yes, but does not significantly affect care	0 (0)	7 (6.5)
Yes, significantly affects care	1 (0.7)	8 (7.4)
Polypharmacy		
<5 prescription medications	83 (58.5)	13 (12.0)
5–8 prescription medications	31 (21.8)	37 (34.3)
>8 prescription medications	28 (19.7)	58 (53.7)
Care organization		
1 main provider	62 (43.7)	31 (28.7)
No main provider or multiple non-coordinated providers	80 (56.3)	64 (59.3)
Significant inter-provider advice confusion	0 (0)	13 (12.0)
Activation in own care		
High	124 (87.3)	53 (49.1)
Moderate	15 (10.6)	32 (29.6)
Low	3 (2.1)	23 (21.3)
Skilled nursing task needs		
None	136 (95.8)	61 (56.5)
Moderate	4 (2.8)	22 (20.4)
High	2 (1.4)	25 (23.2)

ED emergency department, GM general medical inpatient service. Only 140 ED patients out of 142 received SST Global Impression ratings from the observing physician

## Inter-rater Reliability

The physician-physician inter-rater reliability for GI rating computed using Cohen's kappa was 0.60 (SE 0.06) (Table 2). The inter-rater reliability values between service



**Table 2** Cohen's Kappa Scores for Inter-rater Reliability of Global Impression Population Segments

Inter-rater reliability	Number of subjects	Cohen's kappa	Standard error
Service physician-observing physician	135	0.60	0.06
Service physician-nurse	134	0.71	0.06
Observing physician-nurse	136	0.68	0.06

physician-nurse and observing physician-nurse were 0.71 (SE 0.06) and 0.68 (SE 0.06), respectively.

### Predictive Validity

Based on Cox regression, hazard ratios between the highest and lowest medical severity GI categories, adjusted for age and gender, were statistically significant for ED visits (6.94,  $p = 0.012$ ), non-elective hospital admissions (10.81,  $p = 0.004$ ), and mortality (7.44E+10,  $p < 0.001$ ). With the exception of social support, all CFs revealed statistically significant hazard ratios in predicting mortality (Table 3). The Kaplan-Meier survival curves associated with ED visits, hospitalization, and mortality by GI categories are shown in Figures 2a, b, and c, respectively.

## DISCUSSION

In this study, we demonstrated that the SST GI ratings between physicians as well as between physicians and nurses are moderate to substantial<sup>25</sup>. The SST GI and CFs tend to broadly discriminate between individuals with different risks for ED

visits, non-elective hospitalization, and mortality (see Table 3 and Table 4). We observed that individuals rated as category “short decline before dying” experienced a remarkably higher mortality rate compared with individuals in the other categories suggesting that trained clinicians are quite good in making this assessment. Notably, some of the GI categories—categories designed to represent features that reflect distinct clinical needs—show similar rates of adverse medical outcomes. This is not itself an undesirable feature of a needs-based classification scheme. Unlike risk-based classification, the priority here is not to identify individuals with similar risk but rather individuals with similar response to specific actions. While different individuals may have similar risk for poor outcomes and similar needs to reduce those risks, many individuals with similar risk for poor outcomes will have quite different needs. For example, two patients may be readmitted to the hospital within 3 months; however, for one patient, the cause for the readmission was lack of adequate wound care while for the other, the reason was failure to closely monitor and rapidly treat a fragile medical condition. Because the SST captures basic features related to care needs, when collected systematically, it can inform healthcare policy makers regarding the nature and magnitude of health and social services required for their population.

This is the first known validation study of a clinician-administered, healthcare needs-based population segmentation tool in a clinical setting. Commonly, healthcare needs-based population segmentation is performed using an EMR retrospectively even though such records may not always capture information in a reliable or accurate manner.<sup>26</sup> A previous analysis which assessed the reliability of the EMR for administering the SST also found that due to high rates of

**Table 3** Predictive Validity of SST Variables

SST variable ( $n = 250$ , unless otherwise specified)	Cox regression of time to emergency department visit or mortality in 90 days, controlled for age and gender, censored for mortality				Cox regression of time to non-elective hospital admission or mortality in 90 days, controlled for age and gender, censored for mortality				Cox regression of time to mortality in 90 days, controlled for age and gender			
	Hazard ratio	<i>P</i> value	CI-LB	CI-UB	Hazard ratio	<i>P</i> value	CI-LB	CI-UB	Hazard ratio	<i>P</i> value	CI-LB	CI-UB
Global Impression ( $n = 248$ )	6.94	0.012	1.52	31.60	10.81	0.004	2.14	54.52	7.44E+10	< 0.001	8.38E+09	6.61E+11
CF: Function ( $n = 249$ )	2.74	< 0.001	1.57	4.79	2.01	0.017	1.13	3.57	6.70	0.001	2.24	20.05
CF: Social support	2.22	0.012	1.20	4.12	2.07	0.02	1.12	3.83	1.58	0.570	0.32	7.74
CF: Hospital admission	3.74	< 0.001	1.94	7.21	4.19	< 0.001	2.22	7.91	6.03	0.004	1.80	20.21
CF: Disruptive behavior	3.95	0.001	1.70	9.20	3.37	0.005	1.45	7.83	4.96	0.038	1.09	22.53
CF: Polypharmacy	2.69	0.001	1.47	4.92	3.24	< 0.001	1.79	5.87	6.14	0.022	1.30	29.00
CF: Organization of care	4.49	< 0.001	1.97	10.23	3.45	0.002	1.55	7.68	8.90	0.001	2.52	31.44
CF: Activation in own care	2.74	0.003	1.42	5.30	2.64	0.003	1.38	5.03	4.01	0.036	1.09	14.70
CF: Skilled nursing	2.29	0.013	1.20	4.38	1.94	0.05	1.00	3.77	10.28	< 0.001	3.04	34.78

CF complicating factor. Hazard ratios are between highest and lowest complexity levels of all SST variables

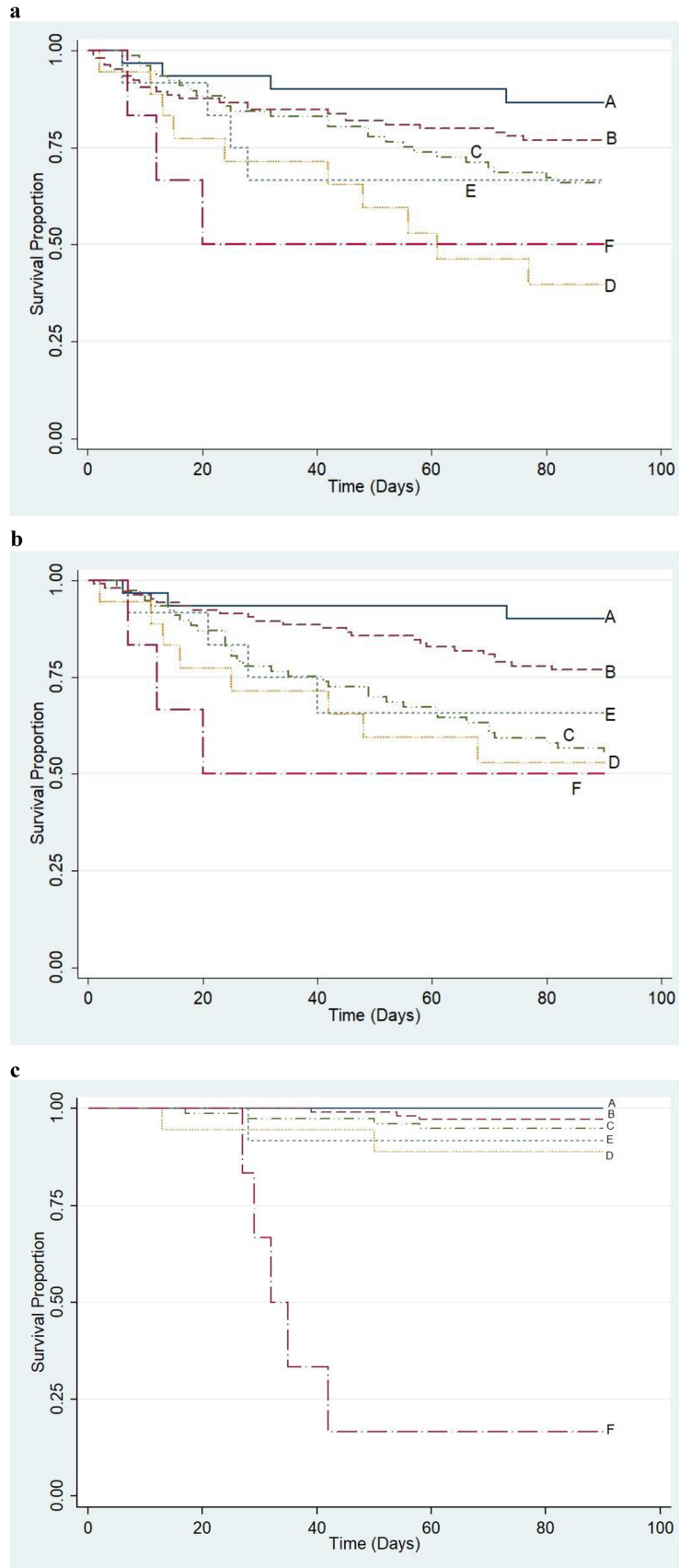


Figure 2 a Kaplan-Meier survival estimates for ED visit outcome by Global Impression categories ( $n = 248$ ). b Kaplan-Meier survival estimates for non-elective hospital admission outcome by Global Impression categories ( $n = 248$ ). c Kaplan-Meier survival estimates for mortality outcome by Global Impression categories ( $n = 248$ ). A, healthy; B, chronic condition, asymptomatic; C, chronic condition, symptomatic; D, long course of decline; E, limited reserve; F, short decline before dying.

**Table 4 Hazard Rate of SST GI Variables, Controlled for Age and Gender**

	Hazard rate of emergency department visit, censored for mortality in 90 days	Hazard rate of non-elective hospital admission, censored for mortality in 90 days	Hazard rate of mortality in 90 days
Healthy	1 [1]	1 [1]	1 [1]
Chronic condition, asymptomatic	1.787 [0.618, 5.167]	2.395 [0.719, 7.973]	1.197e+09*** [1.186e+08, 1.207e+10]
Chronic condition, symptomatic	2.630 [0.902, 7.669]	5.003** [1.512, 16.55]	2.269e+09*** [2.507e+07, 2.054e+10]
Long course of decline	5.300*** [1.606, 17.49]	5.986** [1.547, 23.17]	5.061e+09*** [4.587e+08, 5.585e+10]
Limited reserve	2.688 [0.641, 11.28]	3.903 [0.840, 18.14]	3.796e+09 [3.796e+09, 3.796e+09]
Short decline before dying	6.938* [1.523, 31.60]	10.81** [2.145, 54.52]	7.442e+10*** [8.384e+09, 6.606e+11]
Female gender	1 [1]	1 [1]	1 [1]
Male gender	0.859 [0.536, 1.377]	0.947 [0.598, 1.498]	0.885 [0.315, 2.484]
Age	1.009 [0.985, 1.034]	1.006 [0.982, 1.030]	0.999 [0.949, 1.052]
Observations	248	248	248

Exponentiated coefficients; 95% confidence intervals in brackets  
 \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$

missing information, inter-rater agreement for GI ratings for the same patients between physicians using the EMR and physicians making the assessment based on a face-to-face encounter in the clinic was low ( $\kappa = 0.37$ ).<sup>16</sup>

It is notable that even in the challenging environment of the ED, inter-rater reliability of the SST was good. The finding of substantial inter-rater reliability between physicians and nurses also supports the notion that trained nurses can administer the instrument reliably, increasing the range of sites in which the SST can be implemented practically.

The current work was focused on assessing the inter-rater reliability of elements of the SST and the predictive validity of both GI and CFs as a prerequisite to further application of the instrument. Work is ongoing to identify combinations of SST features that constitute a core group of needs segments based on typical sets of health and health-related social service needs. The goal is to identify unmet needs, the relationship between persistence of unmet needs and outcomes, and strategies for meeting those needs that would be both effective and sustainable at the clinical and organizational level. While we do not expect the SST to simply save costs under the current “model of care,” it may nonetheless promote cost savings when it is used as a key metric for health service innovation (see Appendix for implementation example).

A major strength of this study lies in the strong methodology used for testing inter-rater reliability. We provided all raters of the SST identical stimuli consisting of exposure to a real patient-doctor interaction within an outpatient setting (ED). Second, the raters were working clinicians; demonstrating reliability and validity of SST ratings in this context increases confidence in the feasibility of using the SST in real-world clinical environments. This introduces the possibility that basic healthcare needs of a patient population can be monitored in a simple, ongoing way using the SST and this

information could be useful in monitoring appropriate service levels at the organizational level.<sup>9</sup> Third, by recruiting from the inpatient service as well as the ED, we were able to demonstrate the predictive validity of SST features for a diverse population with a wide range and severity of health conditions.

Regrettably, the evaluation of inter-rater reliability in the ED was restricted to individuals triaged as non-urgent; this limited the number of individuals with more complex medical needs and CFs. Because of the complexity of the inter-rater reliability assessment, we were unable to evaluate inter-rater reliability amongst clinicians on the GM inpatient service. Nonetheless, predictive validity was high across the broad population that included good representation of patients with more severe GI categories and CFs, suggesting that even if the instrument was somewhat less reliable in more complex patients, it retained predictive validity. A second potential limitation is that the population is older and predominantly Chinese, thus primarily speak one of several local languages and dialects rather than English, which our raters are uniformly proficient in.

There are several areas for improvement in this study which can be explored in future studies. They include the measurement of time taken for completion of SST by raters, assessment of raters’ understanding of SST administration, having clinicians and allied health providers complete different parts of the SST, evaluating characteristics of subjects who decline participation, and assessment of how using the SST affects healthcare utilization and outcomes (see Appendix for further discussion on areas for improvement).

In all, our study indicates that clinicians can use the SST to identify patient features related to patterns of clinical healthcare needs as well as complicating factors indicative of broader health and health-related social service needs. Such a

simple, focused assessment can complement the EMR which, even when available, is not a reliable source for key information driving health service needs. A natural extension of this work is to identify packages of health and social services tailored to SST-defined population segments, followed by evaluations of whether receiving population segment-specific service packages improves health outcomes. This work will enable clinical and healthcare policy makers to more effectively gather actionable healthcare needs information from patients, which in turn would facilitate improved allocation of clinical and social services tailored to patient needs.

**Corresponding Author:** David Bruce Matchar, MD; Program in Health Services and Systems Research, Duke-NUS Medical School Singapore, Singapore (e-mail: david.matchar@duke-nus.edu.sg).

**Funding Information** The study at ED was supported by Health Engagement and Action Laboratories (HEALs) from SingHealth Health Services Research Centre (HSRC).

**Data Availability** The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

#### Compliance with Ethical Standards:

Both the ED and GM components of this study were approved by the Singapore SingHealth Centralised Institutional Review Board (CIRB) (CIRB/2016/2005 and CIRB/2016/2629 respectively). Written informed consent to participate and to allow access to the patients' EMRs was obtained from participants or their legal representatives.

**Conflict of Interest:** The authors declare that they do not have a conflict of interest.

## REFERENCES

- US Department of Health and Human Services. Global Health and Aging. World Health Organization; 2011.
- Berwick DM, Hackbarth AD.** Eliminating Waste in US Health Care. *J Am Med Assoc.* 2012;307(14):1513-6.
- Joynt KE, Gawande AA, Orav EJ, Jha AK.** Contribution of Preventable Acute Care Spending to Total Spending for High-Cost Medicare Patients. *J Am Med Assoc.* 2013;309(24):2572-8.
- Stevens A, Gillam S.** Needs assessment: from theory to practice. *Br Med J.* 1998;316:1448-52.
- Steinbach R.** Concepts of Need and Social Justice: Health Knowledge UK; 2009. Available from: <http://www.healthknowledge.org.uk/public-health-textbook/medical-sociology-policy-economics/4c-equality-equity-policy/concepts-need-justice>.
- Higginson IJ, Von Gunten CF.** Population-based needs assessment for patient and family care. In: *Textbook of Palliative Medicine*. Florida: Taylor & Francis Group; 2006:257-264.
- Chong JL, Matchar DB.** Benefits of population segmentation analysis for developing health policy to promote patient-centred care. *Ann Acad Med Singap.* 2017;46(7):287-9.
- Stange KC.** The problem of fragmentation and the need for integrative solutions. *Ann Fam Med.* 2009;7(2):100-3.
- Vuik SI, Mayer EK, Darzi A.** Patient segmentation analysis offers significant benefits for integrated care and support. *Health Aff.* 2016;35(5):765-79.
- Lynn J, Straube BM, Bell KM, Jencks SF, Kambic RT.** Using population segmentation to provide better health care for all: the "Bridges to Health" model. *Milbank Q.* 2007;85(2):185-208.
- Hartgerink JM, Cramm JM, van Wijngaarden JDH, Bakker TJEM, Zorggroep A, Mackenbach JP, et al.** A framework for understanding outcomes of integrated care programs for the hospitalised elderly. *Int J Integr Care.* 2013;13:1-13.
- Baicker K, Chandra A.** Medicare Spending, The Physician Workforce, and Beneficiaries' Quality Of Care. *Health Aff.* 2004;Web Exclusive 4:184-97.
- Shi L.** The Impact of Primary Care: A Focused Review. *Scientifica.* 2012;2012:432892:1-22.
- Chong JL, Lim KK, Matchar DB.** Population segmentation based on healthcare needs: a systematic review. *Syst Rev.* 2019;8(1):202.
- Zhou K, Vidyarthi AR, Wong CH, Matchar D.** Where to go if not the hospital? Reviewing geriatric bed utilization in an acute care hospital in Singapore: Geriatric bed use in an acute hospital. *Geriatr Gerontol Int.* 2017;17(10):1575-1583.
- Chong JL, Low LL, Chan DYL, Shen Y, Thin TN, Ong MEH, et al.** Can we understand population healthcare needs using electronic medical records? *Singap Med J.* 2019;60(9):446-453.
- Buurman BM, Parlevliet JL, Van Deelen JL, De Haan RJ, De Rooij SE.** A randomised clinical trial on a comprehensive geriatric assessment and intensive home follow-up after hospital discharge: the transitional care bridge. *BMC Health Serv Res.* 2010;10:296.
- Chu LW, Pei CK.** Risk factors for early emergency hospital readmission in elderly medical patients. *Gerontology.* 1999;45(4):220-226.
- Glanz K, Rimer BK, Viswanath K.** Health behavior and health education: Theory, Research, and Practice (4th ed.). California: Jossey-Bass; 2008.
- Ahn H, Horgas AL.** Disruptive Behaviors in Nursing Home Residents with Dementia: Management Approaches. *J Clin Outcomes Manag.* 2013;20(12):566-576.
- Cameli D, Francis M, Francois V, Medder N, Von Eden L, Truglio-Londrigan M.** A systematic review of medication reconciliation strategies to reduce medication errors in community dwelling older adults. *JBHI Database Syst Rev.* 2012;10(42 Suppl):1-18.
- AHRQ. National Healthcare Quality Report Rockville, MD: U.S. Department of Health and Human Services; 2010 [cited US Department of Health and Human Services. Available from: <https://archive.ahrq.gov/research/findings/nhqrdr/nhqr10/nhqr10.pdf>.
- Mosen D, Schmittiel J, Hibbard J, Sobel D, Remmers C, Bellows J.** Is Patient Activation associated with Outcomes of Care for Adults with Chronic Conditions? *J Ambul Care Manage.* 2007;30(1):21-29.
- SingHealth. Singhealth Group overall key figures and statistics. 2018. Available from: <https://www.singhealth.com.sg/about-singhealth/newsroom/Documents/SingHealth-AR17-Insert.pdf>.
- Landis JR, Koch GG.** The Measurement of Observer Agreement for Categorical Data. *Biometrics.* 1977;33(1):159-74.
- Chan KS, Fowles JB, Weiner JP.** Review: electronic health records and the reliability and validity of quality measures: a review of the literature. *Med Care Res Rev.* 2010;67(5):503-27.

**Publisher's Note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.