A Global Measure of Patient-reported Outcomes After Injury – Life Back on Track

Gang Chen
Monash Business School Centre for Health Economics
Maame Esi Woode
Monash Business School Centre for Health Economics
Kah Ling Sia
Monash Business School CHE: Monash Business School Centre for Health Economics
Nina Ellis
TAC: Transport Accident Commission
Cassie Citroen
TAC: Transport Accident Commission
Anthony Harris (anthony.harris@monash.edu)
Centre for Health Economics, Monash University https://orcid.org/0000-0003-1641-3320

Research

Keywords: Life Back on Track (LBoT), Transport Accident Commission Longitudinal Study, reasonable predictor of future work status, tracking measure

Posted Date: January 4th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-137181/v1

Version of Record: A version of this preprint was published at Disability and Rehabilitation on January 22nd, 2022. See the published version at https://doi.org/10.1080/09638288.2022.2029958.

Abstract

Background

The study assessed the psychometric properties of the Life Back on Track (LBoT) measure, a new self-reported single-item global measure of the trajectory of wellbeing after a transport accident.

Methods

The data come from four waves of the Transport Accident Commission Longitudinal Study (n=1,556 in wave 1), and two repeated cross-sectional surveys– a Client Outcome Survey (n=5,238) and a Client Experience Survey (n=1,964) of individuals injured in a transport accident in Victoria. The conceptual basis of the measure was confirmed in a qualitative analysis of open-ended survey responses. The psychometric performance of the measure, including known-groups validity, test-retest reliability, sensitivity, and responsiveness was investigated.

Results

The LBoT measure was found to have a conceptual basis of recovery towards a normal life with domains of independence, control, happiness, work, social life, pain, physical function, cognitive function, work and leisure activities, income, anxiety, and depression. There were significant differences in the distribution of the LBoT scores based on the respondent's depression, pain, return to work status, financial ability to get by, ability to cope, and ability to bounce back. The LBoT measure was a reasonable predictor of future work status and was moderately responsive to change.

Conclusions

LBoT is a valid measure to track the individual's trajectory of subjective wellbeing in the context of recovery after a trauma, and it covers wider concepts than health-related quality of life. For use as a performance or tracking measure, however, further evidence is needed on its responsiveness.

Introduction

Transport accidents remain a serious and complex public health issue globally (World Health Organization, 2018). In many developed countries, the burden has shifted from mortality towards injury and disability in the last 20 years. Measuring the burden of injury at the individual level is important if we are to make rational decisions about interventions and policies that reduce injury and improve rehabilitation.

There is no agreement on how best to measure and track the overall quality of life (QoL) or subjective wellbeing of transport accident survivors (Rissanen, et al., 2017). Most multi-dimensional health related quality of life (HRQoL) measures focus on health-related domains (physical and mental health components), although the 20-item QoL questionnaire covers domains related to satisfaction, perception,

health and social, and economics (Chaikoolvatana & Sripech, 2007). There is, however, a clear demand to develop and use a valid single (non-monetary) metric to integrate changes in years of life with changes in disability, functionality, pain, and emotional distress to measure the impact of injury (Segui-Gomez & MacKenzie, 2003) and potentially other kinds of recovery from trauma.

Following extensive discussion with clients, the Victorian Transport Accident Commission (TAC) and the Social Research Centre at the Australian National University devised a single-item global scale, Life Back on Track (LBoT), to measure individual self-appraisal of recovery following a transport accident. It asks a simple single question “How would you rate the extent to which you have been able to 'get your life back on track', on a scale from 1 to 10?”

As a direct global measure of subjective personal wellbeing the LBoT measure has similarities to measures of life satisfaction or happiness such as the single domain Satisfaction with Life Scale (SWLS) (Pavot & Diener, 2008), and single-item global life satisfaction measure used in large surveys (e.g. Household, Income and Labour Dynamics in Australia survey; German Socio-Economic Panel) and its conceptual basis has much in common with the concept of recovery in the mental health literature (Anthony, 1993).

Like other subjective wellbeing measures, LBoT allows a within-person QoL construct dynamism, as what matters to the individual and how much it matters (the standard by which they judge QoL) can change over time (Allison, et al., 1997). To the extent that the LBoT measure has been designed explicitly in the context of a trajectory of recovery with a pre-injury reference point, the standard by which an individual judges wellbeing may be more stable than in other measures of subjective wellbeing (SWB). There are a number of multiple domain SWB measures, but only those with a small number of items are coherent (Cummins & Weinberg, 2015). There is some evidence that single-item global life satisfaction measures perform very similarly compared to the multiple-item SWLS (Cheung & Lucas, 2014); and in health related quality of life there is evidence that a global rating showed satisfactory construct validity in terms of measurement of health perception and physical functioning (de Boer, et al., 2004). It remains unclear however if a global subjective wellbeing measure is sufficiently reliable or responsive to be useful in program or performance evaluation.

The original purpose of the LBoT measure is to provide a global performance metric for services and support to those who are injured in an accident that captures the affected components of wellbeing, including not only mental and physical health but also work, social, emotional, and financial aspects of QoL. This study evaluated the psychometric performance of the LBoT measure (criterion and construct validity, reliability, sensitivity, and responsiveness) in a sample of people who have made an injury claim to a universal public transport injury insurer in the state of Victoria (TAC) following a transport accident.

Methods

Data

The assessment of psychometric properties of the LBoT measure is based on 3 telephone surveys of Victorians with a managed TAC claim for injury in a transport accident: four waves of a Longitudinal Study survey 2012-2016 with 1,556 respondents in wave 1; and two repeated cross sectional surveys conducted on behalf of the TAC from October 2011-2017, the Client Outcome Survey (COS) with 5,238 respondents and the Client Experience Survey (CES) with 1,964 interviews from 604 unique respondents. While the Longitudinal and COS surveys focused on recovery outcomes, the CES survey focussed on perceptions of service and included respondents who had more enduring disabilities. The three surveys are described in detail in Electronic Supplementary Material 1.

The Life Back on Track (LBoT) Measure

Respondents were asked to rate whether they considered their lives to be back on track as:

“In other research, TAC clients often talk about trying to 'GET THEIR LIFE BACK ON TRACK' following a transport accident. This can mean different things to different people. Thinking about your own circumstances right now (today), how would you rate the extent to which you have been able to 'get your life back on track', on a scale from 1 to 10, where 1 means ‘not at all’, and 10 means ‘completely back on track’?” In the CES for respondents who had more enduring disabilities, the recall time was 2 weeks preceding the interview.

Analyses

Construct Validity

Construct validity is the degree to which the LBoT measure captures what it intends to measure – in this case general subjective wellbeing in recovery from a transport accident. A known-groups validation was conducted, based on the principle that certain specified groups of TAC clients are expected to score LBoT differently from others, and the LBoT measure should be sensitive to these differences.

The known-groups were identified based on a construct analysis of qualitative survey data (see Supplementary Material 2) and represented by the following indicators: (a) self-reported injury severity levels (a 5-point scale: Very Severe, Severe, Moderate, Slight and Very Slight); (b) depression subscale of the Depression Anxiety Stress Scales (DASS)-21 (Lovibond & Lovibond, 1995; Henry & Crawford, 2005) (which contains 7 items, each with a four-point severity scale: none of the time, some of the time, a lot of the time, most of the time); (c) pain; (d) financial ability to get by (a 4-point scale: with great difficulty, with some difficulty, fairy easily, very easily); (e) expected time to recovery (4 response levels: already recovered as much as possible, will be in the next few months or so, will be within a year, and will take longer than a year); (f) ability to cope with their injuries given its nature (a 5-point Likert scale: very poor, poor, moderate, good, and very good) (Lazarus, 1993; Hepp, et al., 2011), and (g) ability to bounce back from the accident (on a 10-point scale ranging from strongly disagree to strongly agree).

Pain was measured using the Numerical Rating Scale (NRS) in the TAC Longitudinal Study. This validated scale asks respondents to rate their level of pain on a scale of 0 to 10, where 0 is no pain at all and 10 worst possible pain (Sendbeck, et al., 2015; Thong, et al., 2018). The rating was then recoded as none (0), mild (1–2), moderate (3–5), strong (6–8), severe (9–10). In the COS, respondents were asked "the amount of bodily pain they had in the past 7 days" and had to choose between 6 options: 1 – none, 2 – very mild, 3 – mild, 4 – moderate, 5 – severe, and 6 – very severe.

The Kruskal-Wallis H test was used to test for statistically significant overall differences between groups, and Dunn's test for differences in pairs between groups (presented in the Supplementary Material Table 3). A p-value <0.05 was taken as statistically significant in this and all other hypotheses tests in the paper.

Criterion Validity

Criterion validity, in this context, refers to the extent to which the LBoT measure correlates with an external standard measure. Criterion validity is commonly assessed through the investigation of the concurrent validity and predictive validity of the measurement. Concurrent validity involves comparing the LBoT measure to a standard measure of wellbeing in recovery at the same time point. There is no gold standard for wellbeing in recovery and based on the availability of data we focused on a closely related conept of HRQoL as measured by the EQ-5D-3L - the most widely used preference-based health-related quality of life instrument in the world (Richardson, et al., 2014). The EQ-5D-3L contains five dimensions with each dimension measured using one item (which includes mobility, usual activities, self-care, pain/discomfort, and anxiety/depression) as well as a single-item EQ-VAS (EuroQoL Group, 1990). Concurrent validity was measured using the Spearman correlation coefficients between LBoT scores and overall HRQoL scores from the EQ-5D-3L utility scores and the EQ-VAS (visual analogue scale). All five dimensions were assessed using three response levels (no problem, some problems, and extreme problems). The EQ-VAS lies on a scale of 0 (worst imaginable health state) to 100 (best imaginable health state). The EQ-5D-3L asked respondents about their health 'today' (i.e. at the day of the interview). The EQ-5D-3L was scored by using the original UK tariff (Dolan, 1997).

Predictive validity was assessed using LBoT scores in wave 1 of the Longitudinal Study in a logit model predicting return to the same job with the same duties and employer, conditional on being in employment prior to the accident and controlling for injury severity, age, gender, education, employment type and country of birth. Validity was assessed as the ability of the logit regression to classify post accident employment status correctly using the c statistic (area under the receiver operating characteristic curve). A value closer to 1 and further from 0.5 suggests greater discrimination and therefore stronger validity (Hosmer & Lemeshow, 2000).

Reliability

Reliability refers to the extent to which LBoT scores are affected by random error. We focus on whether LBoT is consistent across time. The Longitudinal Study was used to identify clients who were in a stable condition across two survey time periods (waves 1 and 2). As the time interval between the two waves was relatively long (i.e. ≥ 3 months), we constructed samples with TAC clients who were in a relatively stable condition between waves as indicated by individual and combinations of scores on measures of pain, financial ability to get by, DASS score, a single-item global health rating from the Short-Form 12 Health Survey (SF-12) (Ware, et al., 1996), main labour market activity, and vocational status.

Patients were defined as being in a stable condition if they gave the same pain rating, the same DASS group (Normal, Mild, Moderate, Severe, Extremely Severe), same SF-12 rating and same vocational status in waves 1 and 2. Reliability was measured by the intraclass correlation coefficient (ICC). An ICC of 0.75–0.90 is generally classified as good while an ICC larger than 0.9 is classified as high or excellent reliability (Koo & Li, 2016).

Sensitivity and Responsiveness

Sensitivity of a measure is the ability to detect differences between groups while responsiveness is the ability to detect changes. Sensitivity was evaluated as the extent of the difference in response from those who reported that they had recovered to those who said they had not. We chose pain levels and employment status to measure responsiveness, that is, respondents who were considered to be in a stable condition across two waves (same pain rating and same vocational status) were excluded. To adjust for potential bias due to non-response, scores were weighted by the inverse probability of non-response from wave 1 to each of the subsequent waves. Two sets of inverse probability weights were used, one based on the average non-response to that time, and the other based on the probability of non-response for all three periods after wave 1 (see Supplementary Table 3.4). The probability of non-response was estimated with a logit model with age, gender, education, area of residence (rural vs metro), longitudinal survey cohort, involvement in accident (road user), injury type, injury severity, recovery expectation, language spoken at home and country of birth as covariates.

To measure both sensitivity and responsiveness, the effect size and the standard response mean was used. A minimum effect size of 0.41 is recommended with an effect size of 1.15 considered as moderate and 2.70 considered as strong (Ferguson, 2009). A standard response mean of 0.5 is generally considered as indicating moderate responsiveness with a value of 0.8 and above indicating strong responsiveness (Ferguson, 2009).

The existence of ceiling and/or floor effect can threaten responsiveness and cause measurement inaccuracy. The potential ceiling and floor effects of the LBoT measure were examined based on the full sample as well as by gender and age groups. Ceiling or floor effects are taken as evident if ≥ 15% respondents scored the best or the worst of a measure (McHorney & Tarlov, 1995).

Results

Sample characteristics

The socio-demographic characteristics of the respondents in wave 1 of the Longitudinal Study are presented in Table 1. The age distribution of the respondents was roughly even, while all three surveys had more males than females. A little over 75% of the respondents had less than a bachelor's degree across all datasets. A little over half of the respondents (54%) were drivers of a vehicle during the accident, with slightly less than a fifth (18%) being motorcyclists during the accident. Respondent characterisics in the other two surveys are similar (see Supplementary Material 1).

Table 1: Respondents' socio-demographic characteristics

Longitudinal Study survey (wave 1)		
Variable	N	%
Age		
15-24	278	17.8
25-34	289	18.5
35-44	277	17.7
45-54	295	18.8
55-64	211	13.5
65+	216	13.8
Gender		
Female	659	42.1
Male	907	57.9
Road User		
Cyclist	129	8.2
Driver	840	53.6
Motor Cyclist	284	18.1
Passenger	162	10.3
Pedestrian	139	8.9
Train/Tram/Unknown	12	0.8
Education		
Bachelor degree or higher	347	22.3
Diploma/certificate	588	37.8
Primary & Decondary School	615	39.6
Other	5	0.3
Residence		
Metro	1,013	64.7
Rural	553	35.3
Country of Birth		
Overseas	438	28.0

Australia	1,125	72.0

Psychometric properties of the LBoT measure

Descriptive statistics of the LBoT measure

The mean, standard deviation (SD), and the distributions of the LBoT scores across the different time-periods and across the three surveys are presented in Figure 1. The mean of the LBoT scores was higher in the COS and wave 1 of the Longitudinal Study compared to respondents in the CES. The mean LBoT score increased over time with much of the change occurring by the second wave (3 months later) when the median increased from 7 to 8. There is an increase in the proportion of respondents with a score of 10, but given the 41% loss to follow-up to wave 4 this trend should be interpreted with caution.

Known-Groups Validity

We expected that respondents who reported more severe injuries, depression, or pain would report significantly lower LBoT scores, while those who reported having better ability to cope with their injuries and to bounce back from their injuries would report higher LBoT scores. These hypotheses were supported by the Kruskal-Wallis H test statistics (p-value < 0.001) and the results are robust over time (wave 2 to wave 4) as shown in Table 2.

Table 2: Known-Groups Validity: LBoT score by group

	Categories	N	Mean	SD	p-value
Injury Severity	Very Severe	296	5.4	2.7	<0.001
	Severe	590	6.4	2.5	
	Moderate	465	7.5	2.2	
	Slight	105	8.4	2.1	
	Very Slight	41	8.6	2.0	
General Health (Wave 1 vs Pre)	Worse Health	885	6.0	2.5	<0.001
	Same Health	521	7.8	2.2	
Depression (DASS)	Normal	1023	7.4	2.3	<0.001
	Mild	114	5.7	2.2	
	Moderate	127	4.8	2.3	
	Severe	42	4.4	1.9	
	Extremely Severe	69	3.7	2.5	
Pain	None (0)	392	8.4	2.1	<0.001
	Mild (1-2)	118	7.9	2.1	
	Moderate (3-5)	436	6.7	2.3	
	Strong (6-8)	500	5.5	2.3	
	Severe (9-10)	65	4.2	2.5	
Time to Recovery	Already Recovered	244	8.8	1.8	<0.001
	Next Few Months	569	7.2	2.2	
	Within a Year	319	6.0	2.2	
	Longer than a Year	92	5.4	2.4	
	Will Not Recover	32	3.7	2.8	
Return to Work	Paid Work, Same Employer	617	7.8	2.0	<0.001
	Paid Work, Different Employer	49	7.4	2.4	
	Retired/Permanently Unable to Work	247	6.2	2.8	
	Not Working, Pre & Post	171	5.9	2.6	
	Paid Work, Pre	390	5.5	2.5	

Financial Ability to Get By	Same Difficulty	887	7.0	2.4	<0.001
	Worse Difficulty	375	5.7	2.6	

Table 2(cont.): Known-Groups Validity

	Categories	N	Mean	SD	p-value
Ability to Cope	Very poor	60	5.0	2.8	<0.001
	Poor	169	4.8	2.5	
	Moderate	483	6.3	2.3	
	Good	527	7.3	2.3	
	Very good	254	8.2	2.2	
Ability to Bounce Back (Recoded): Difficult to Bounce Back	1= Strong Disagreement	484	8.4	2.0	<0.001
	2	164	7.1	2.0	
	3	284	6.6	2.0	
	4	281	5.8	2.1	
	5= Strong Agreement	265	4.6	2.6	

Kruskal-Wallis H Test p-value

Concurrent Validity

Table 3 shows that there was moderate correlation between the LBoT measure and the EQ-5D-3L at all time-periods (TAC Longitudinal Study, wave 1 to wave 4) and for the COS data. The magnitude of correlations was stronger between LBoT and EQ-5D-3L utility scores (r = 0.671 to 0.732) than with EQ-VAS (r = 0.567 to 0.647), across all surveys.

Table3: Concurrent Validity Spearman Rank Correlation

				Spearman Rank Correlation	
	N	Mean	SD	Coefficient	p-value
T1					
EQ-VAS	1506	71.61	20.77	0.57	<0.001
EQ-5D-3L Utility	1488	0.62	0.32	0.67	<0.001
T2					
EQ-VAS	1208	75.89	19.81	0.64	<0.001
EQ-5D-3L Utility	1189	0.68	0.30	0.70	<0.001
T3					
EQ-VAS	1063	76.83	19.19	0.65	<0.001
EQ-5D-3L Utility	1049	0.70	0.30	0.73	<0.001
T4					
EQ-VAS	833	0.72	0.30	0.65	<0.001
EQ-5D-3L Utility	833	0.75	0.25	0.70	<0.001
COS					
EQ-VAS	5348	0.67	0.32	0.64	<0.001
EQ-5D-3L Utility	5438	70.99	21.38	0.69	<0.001

Predictive Validity

LBoT scores at wave 1 were predictive of return to work at each subequent time period (Table 4). Individuals with a higher LBoT score were generally more likely to have returned to work in later periods. All coefficients are statistically significant (all p<0.001), with AUC values of 0.73 in wave 2 to almost 0.8 in wave 4 suggesting the LBoT measure is a reasonable predictor of future work status post injury.

Table 4: Predictive Vaidity Return to Work

		N	Model	Coefficient
LBoT at wave 1	At wave 2	670	OR	1.52***
				(0.08)
			AUC	0.74
	At wave 3	509	OR	1.61***
				(0.15)
			AUC	0.76
	At wave 4	330	OR	1.87***
				(0.27)
			AUC	0.79

Notes: A LOGIT model was used. In addition to what have been reported in the table, other control variables include respondents' perceived injury severity, injury type, age, gender, education, employment at time of accident, type of household, country of birth, whether they live in the metro area or a rural area, their role in the accident (driver, passenger, motorcyclist etc.) and the language they speak at home.

Odds ratio (OR) reported. Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, AUC: Area under the Curve.

Test-Retest Reliability

In the 3 months between waves 1 and 2 respondents who had stable response measures of pain, financial ability to get by, DASS score, global health rating, main labour market activity, and vocational status reported consistent LBoT scores, with an ICC of at least 0.76 across all combinations tested.

Sensitivity

Those who say that they had recovered gave a higher LBoT score than those who said they had not recovered. The absolute value of the effect size was 1.538-2.461 across the waves of the Longitudinal Study and 1.523 in the COS data. These results indicate that, irrespective of the sample used, the LBoT measure was able to detect different recovery status with moderate effect.

Responsiveness

Overall, we observed a ceiling effect for LBoT scores for the three different surveys, across age groups and gender. The percentage of respondents who scored 10 was 18-26% across the Longitudinal and COS surveys. Only 4% in the first two sample and 7% of those classified as having a disability reported a score of 0.

Table 5 shows that the estimated effect sizes on returning to work or being in pain between waves were all below a conventional threshold of 0.8, and could only be considered moderately responsive using the SRM criterion if we consider the longer term change from wave 1 to wave 4. Adjusting for potential attrition bias using inverse probability weightings did not affect these results irrespective of the time period used to weight. It is worth noting that, while the LBoT measure did not demonstrate responsiveness to change in this sample, it was at least as responsive as the EQ-5D-3L utility measure (see Supplementary Materials, Table 3.5).

Table 5: Responsiveness of the LBoT measure

	N	Mean	Diff Mean	SD	Diff SD	SD Diff	ES	SRM
LBoT wave 1 vs.	830	6.4	0.7	2.5	0.0	2.2	0.3	0.3
LBoT wave 2	830	7.0		2.5				
LBoT wave 1 vs.	814	6.4	1.0	2.5	-0.1	2.2	0.4	0.5
LBoT wave 3	814	7.4		2.4				
LBoT wave 1 vs.	715	6.4	1.2	2.4	0.0	2.3	0.5	0.5
LBoT wave 4	715	7.6		2.5				

ES: Effect size, SD: Standard deviation, SRM: Standard Response Mean, Diff: Difference; IPW: inverse probability weights

Respondents whose pain level and employment status remained the same between two waves were regarded to be in a stable condition and excluded from this analysis.

Discussion

This study provides empirical evidence on the psychometric properties (validity, reliability, sensitivity and responsiveness) of a novel measure of wellbeing in recovery following a transport accident – Life Back on Track. As a post injury outcome measure, the LBoT has potential to be used for priority setting in injury prevention, treatment and rehabilitation. It is likely to be particularly useful where there is a focus on the long-term health and wider consequences of injury, and the ability of people to adapt and cope with the consequences of physical or emotional trauma.

A qualitative analysis of open-ended text responses in the survey (reported in the electronic supplementary material) suggests that the reported domains of the LBoT measure cover the concepts of recovery towards a normal life in terms of independence, control, happiness, work, social life with family and friends, pain, physical function, cognitive function, work and leisure activities, income, anxiety and depression. The smaller sample of clients who answered the question of what the measure means to them reported a somewhat wider set of concepts including independence and living a normal life, where normal for some was life before the accident and for others appeared to be accepted as an altered life. For these clients the distinction between being “on track” and “back on track” was more evident and brings into focus the performance of the measure as both an indicator of recovery and an indicator of quality of life.

Overall, the empirical evidence indicates that the LBoT measure is a valid measure to track the individual’s trajectory of SWB in recovery after a trauma. We found that the LBoT measure was able to distinguish the pre-defined known groups, such as injury severity and self-reported health status. The reliability of the LBoT measure was judged high by the ICC regardless of the way a stable cohort between the first two waves of the longitudinal data was constructed. The LBoT measure was found to be a moderately sensitive instrument when distinguishing between different self-evaluated recovery status. The psychometric properties of the LBoT measure are as strong as those measured for other single-item quality of life instruments. One single-item global life satisfaction scale reported high criterion and construct validity (Cheung and Lucas 2014), while the linear analogue self-assessment (LASA) for overall quality of life in cancer patients undergoing treatment, showed convergent and discriminant validity at baseline, but the evidence was less clear under treatment (Bernhard et al 2001).

On the other hand, the LBoT measure was not found to be very responsive. However, this result should be interpreted with caution since the longitudinal surveys provided imperfect data on changes in health status or wellbeing over time. It is in the context of the ability of the data to provide a reliable measure of change that we note that the LBoT was at least as responsive as the widely used quality of life measure EQ-5D-3L in this data. There is currently limited evidence on the comparative responsiveness of single-item life satisfaction measures given that few longitudinal studies have included both single and multi-item measures of life satisfaction. In studies of cancer patients, Bernhard et al. (2001) reported that the responsiveness of single-item LASA at 3 months after chemotherapy was comparable with indicators of physical wellbeing and coping, whereas de Boer et al. (2004) reported that for a single global quality of

life measure using the VAS, the distribution-based responsiveness at five weeks post-operatively was moderate (SRM=-0.47; effect size=-0.56) and anchor-based responsiveness to detect clinically relevant changes was reported to show 'good' correlation (r=0.55) compared to multi-item questionnaires.

Moreover, the ceiling effect of the LBoT measure, i.e. the lack of discriminative ability (or precision) towards the highest end of LBoT distribution, may mean that we did not have enough power to detect a response based on the current sample size. It should be noted that the ceiling effect we observed is similar to that found for the EQ-5D-3L (i.e. scored as full health according to the health state classification system) in the same dataset (19% scored 10 in LBoT vs. 16% scored 1 based on EQ-5D-3L).

LBoT is a single item global measure of the current experience of wellbeing, but in contrast to commonly used measures of current life satisfaction or happiness it has both an explicit life event reference point (the accident), and an explicit focus on a personal ideal life trajectory. The ability to *get one's life back on track* following a trauma suggests a measure that encourages consideration of active aspects of satisfaction such as personal participation and personal control. Two people with radically different circumstances pre and post an accident may both reasonably state that their life was equally back on track; conversely two individuals who are returned to the same level of functioning may report very different degrees of being "back on track". It may be that the LBoT measure has elements of a measure based on a basic justice approach to measuring desirable outcomes – evaluating health and social care programs not only in terms of achieving an improvement in reported quality of life, but also in regard to their effects on the capacity to achieve a preferred quality of life (Sen, 1999).

There are two main limitations of this study. First, the term Life Back on Track was initially developed from language used by clients in previous qualitative research studies, and the underlying concept was explored in the current study in a content analysis of responses to simple questions on the meaning of the concept "life back on track". While the concepts reported have face validity, it is not clear that this approach has resulted in a full understanding of the conceptual basis of the measure. Second, the analysis of the responsiveness of the LBoT measure was limited to proxy measures of the recovery status of the respondents; and for those with an enduring disability, although cross section data showed similar results to other clients, an absence of repeated measurement meant that we were unable to test responsiveness adequately.

Conclusion

While further evidence on its responsiveness to change needs to be established, the LBoT measure appears to be a valid measure to track an individual's trajectory of subjective wellbeing at least in the context of recovery after a transport accident. LBoT covers wider concepts than health-related quality of life domains. The context differentiates the measure from standard single item SWB measures (life satisfaction and happiness), and the measure has a conceptual basis in ideas of recovery, physical and social functioning, wellbeing and employment. It is possible that respondents will adapt their attitudes and life aspirations to their changed circumstances and that changes in LBoT scores over time may in

part reflect this. This kind of adaptation is not unique to this context, but the LBoT measure, by focussing on a trajectory anchored on a past ideal view of a good life, may be less susceptible to this source of shifts in response compared to other SWB measures.

Declarations

Funding: The study is funded by the TAC Research Project (T019). We are grateful to the survey participants from the TAC Longitudinal Study, the Client Outcomes Survey (COS), and the Client Experience Survey (CES).

Conflicts of interest/Competing interests: Nina Ellis and Cassie Citroen are employed by the Transport Accident Commission. All other authors have no competing interests.

GC designed the quantitative analysis and contributed to writing the manuscript. MW performed the quantitative analysis and contributed to writing the manuscript. KLS reviewed the literature and contributed to the writing of the manuscript. NE and CS were responsible for the data collection and initial instrument development with clients. AH analysed the qualitative data and contributed to the writing of the manuscript. All authors read and approved the final manuscript.

Availability of data and material (data transparency): Survey data used in the study contain confidential information and not released publicly

Code availability (software application or custom code): Not applicable

Ethical approval: Ethics approval was granted by the Monash University Human Research Ethics Committee, Monash University, Australia (Reference no. 18278). The study adhered to the principles outlined in the Declaration of Helsinki.

Informed consent: Informed consent was obtained from all individual participants included in the study.

References

1. Allison, P. J., Locker, D. & Feine, J. S., 1997. Quality of life: A dynamic construct. *Social Science & Medicine,* 45(2), pp. 221-30.
2. Anthony, W. A. 1993. Recovery from mental illness: The guiding vision of the mental health service system in the 1990s. *Psychosocial Rehabilitation Journal*, 16(4), 11-23. doi:10.1037/h0095655.
3. Australian Institute of Health and Welfare, 2017. *Impact of overweight and obesity as a risk factor for chronic conditions,* Cat. no. BOD 12. Canberra: AIHW.
4. BITRE, 2018. *Road Trauma Australia 2017 Statistical Summary,* Canberra ACT: Bureau of Infrastructure, Transport and Regional Economics.

5. Brazier, J. E., Yang, Y., Tsuchiya, A. & Rowen, D. L., 2010. A Review of Studies Mapping (or cross walking) Non-Preference Based Measures of Health to Generic Preference-Based Measures. *European Journal of Health Economics,* 11(2), pp. 21-225.
6. Chaikoolvatana, A. & Sripech, S., 2007. Quality of Life of Road Accident Head-Injured Patients after Craniotomy. *The Southeast Asian Journal of Tropical Medicine and Public Health,* 38(4), pp. 761-768.
7. Cheung, F. & Lucas, R. E., 2014. Assessing the Validity of Single-item Life Satisfaction Measures: Results from Three Large Samples. *Quality of Life Research,* 23(10), pp. 2809-2818.
8. Cummins, R. A. & Weinberg, M. K., 2015. Subjective Wellbeing – multi-item measurement: A review. In: *Global Handbook of Wellbeing and Quality of Life.* Dordrecht: Springer.
9. de Boer, A. G. E. M. et al., 2004. Is a single-item visual analogue scale as valid, reliable and responsive as multi-item scales in measuring quality of life?. *Quality of Life Research,* Volume 13, pp. 311-320.
10. Dolan, P., 1997. Modeling Valuations for EuroQol Health States. *Medical Care,* 35(11), pp. 1095-1108.
11. Economic Connections, 2017. *Cost of Road Trauma in Australia - Summary Report,* Canberra ACT: Australian Automobile Association.
12. EuroQoL Group, 1990. EuroQoL-a New Facility for the Measurement of Health-Related Quality of Life. *Health Policy,* 16(3), pp. 199-208.
13. Ferguson, C. J., 2009. An Effect Size Primer: A Guide for Clinicians and Researchers. *Professional Psychology: Research and Practice,* 40(5), pp. 532-538.
14. Hanley, J. A. & McNeil, B. J., 1982. The Meaning and Use of the Area Under a Receiver Operating Characteristic (ROC) Curve. *Radiology,* 143(1), pp. 29-36.
15. Henry, J. D. & Crawford, J. R., 2005. The Short-Form Version of the Depression Anxiety Stress Scales (DASS-21): Construct Validity and Normative Data in a Large Non-Clinical Sample. *British Journal of Clinical Psychology,* 44(Pt 2), pp. 227-239.
16. Hepp, U. et al., 2011. The Long-term Prediction of Return to Work following Serious Accidental Injuries: a Follow Up Study. *BMC Psychiatry,* Volume 11, p. 53.
17. Hosmer, D. & Lemeshow, S., 2000. *Applied Logistic Regression.* New York(NY): John Wiley & Sons.
18. Koo, T. K. & Li, M. Y., 2016. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine,* 15(2), pp. 155-163.
19. Lazarus, R., 1993. From Psychological Stress to the Emotions: a History of Changing Outlooks. *Annual Review of Psychology,* Volume 44, pp. 1-21.
20. Lovibond, S. & Lovibond, P., 1995. *Manual for the Depression Anxiety Stress Scales.* Sydney: Psychology Foundation Monograph.
21. Luchman, J. N., 2013. *DOMIN: Stata Module to Conduct Dominance Analysis.* s.l.:Boston College Department of Economics, revised 04 May 2015.
22. McHorney, C. A. & Tarlov, A. R., 1995. Individual-Patient Monitoring in Clinical Practice: Are Available Health Status Surveys Adequate?. *Quality of Life Research,* 4(4), pp. 293-307.

23. Pavot, W. & Diener, E., 2008. The Satisfaction With Life Scale and the emerging construct of life satisfaction. *Journal of Positive Psychology,* 3(2), pp. 137-152.
24. Richardson, J., McKie, J. & Bariola, E., 2014. Multiattribute Utility Instruments and their Use. In: C. A. J., ed. *Encyclopedia of Health Economics.* San Diego, CA: Elsevier Science, pp. 34-357.
25. Rissanen, R., Berg, H.-Y. & Hasselberg, M., 2017. Quality of Life Following Road Traffic Injury: A Systematic Literature Review. *Accident Analysis & Prevention,* Volume 108, pp. 308-320.
26. Scholz, U., Dona, B. G., Sud, S. & Schwarzer, R., 2002. Is General Self-Efficacy a Universal Construct? Psychometric Findings from 25 Countries. *European Journal of Psychological Assessment,* 18(3), pp. 242-251.
27. Schwarzer, R. & Jerusalem, M., 1995. Generalized Self-Efficacy Scale. In: *Measures in Health Psychology: a User's Portfolio. Casal and Control Beliefs.* Winsor, UK: NFER-NELSON, pp. 35-37.
28. Segui-Gomez, M. & MacKenzie, E. J., 2003. Measuring the Public Health Impact of Injuries. *Epidemiologic Reviews,* 25(1), pp. 3-19.
29. Sen, A., 1999. *Development as Freedom.* Oxford: OUP.
30. Sendbeck, M., Araujo, G. G., Schett, G. & Englbrecht, M., 2015. Psychometric Properties of Three Single-Item Pain Scales in Patients with Rheumatoid Arthritis Seen During Routine Clinical Care: a Comparative Perspective on Construct Validity, Reproducility and Internal Responsiveness. *RMD Open,* Volume 1, p. e000140.
31. The Social Research Centre, 2016. *Client Outcomes Research Wave 9 - Methodological Report,* Melbourne: Social Research Centre.
32. The Social Research Centre, 2017. *'Your Journey to Recovery' Transport Accident Commission Longitudinal Study - T5 Technical Report,* Melbourne: Social Research Centre.
33. Thong, I. S. K., Jensen, M. P., Miro, J. & Tan, G., 2018. The Validity of Pain Intensity Measures: What do the NRS, VAS, VRS, and FPS-R Measure?. *Scandinavian Journal of Pain,* 18(1), pp. 99-107.
34. Wallis Market and Social Research, 2017. *Client Experience Survey: June 2017 Benchmark Wave - Technical Report,* Geelong: Transport Accident Commission.
35. Ware, J. E., Kosinski, M. & Keller, S. D., 1996. A 12 Item Sohrt Form Health Survey: COnstruction of Scales and Preliminary Tests of Reliability and Validity. *Medical Care,* Volume 34, pp. 220-233.
36. World Health Organization, 2018. *Global Status Report on Road Safety 2018,* Geneva: WHO.

Figures

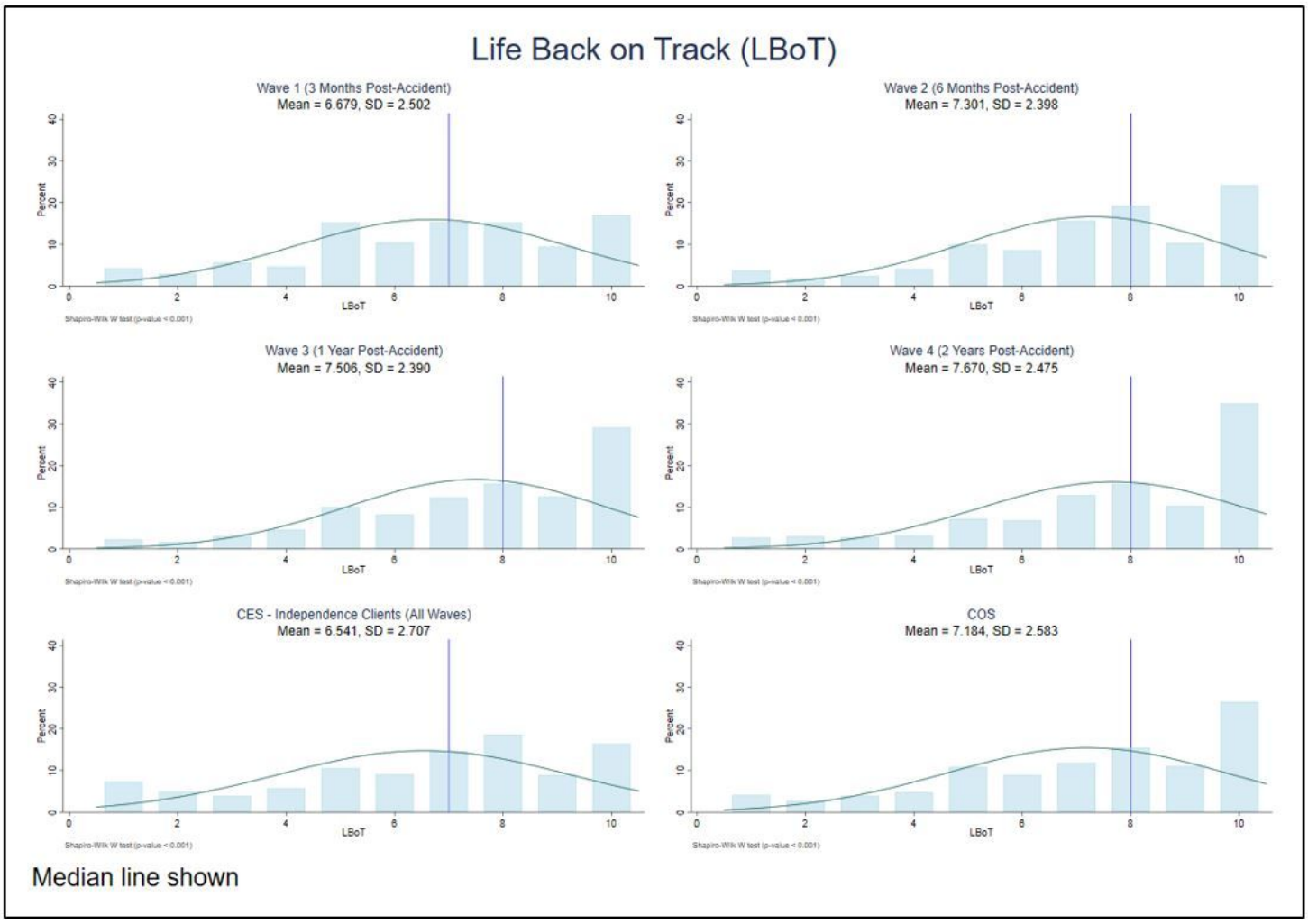

Figure 1

Distribution of Life Back on Track scores across surveys

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- ElectronicSupplementaryMaterial.docx