A Taxonomy to Characterize Stressor Variation in Studies of Physical Resilience and its Validation in Total Knee Replacement

Authors:

Frederick Sieber, MD 4 (fsieber1@jh.edu)

Mallak K Alzahrani, MS 1,2 (malzahr1@jh.edu)

Qian-Li Xue, PHD 1, 2 (qxue1@jhu.edu)

Ravi Varadhan, PHD 6 (ravi.varadhan@jhu.edu)

Thomas Laskow, MD 1,2 (tlaskow1@jhmi.edu)

Charles Brown, MD 4 (cbrownv@jhmi.edu)

Brian Buta 1 (bbuta@jhu.edu)

Julius Oni, MD 7 (joni1@jhmi.edu)

Fangyu Liu 1,2 (fliu28@jhmi.edu)

Jeremy Walston, MD*1 (jwalston@jhmi.edu)

Karen Bandeen-Roche, PHD*5 (kbandee1@jhu.edu)

Affiliations:

(1) Department of Medicine Division of Geriatric Medicine and Gerontology, School of Medicine, Johns Hopkins University, Baltimore, MD, USA

(2) Center on Aging and Health, Johns Hopkins Medical Institutions, Baltimore, MD, USA

(3) Department of Medicine Division of General Internal Medicine, School of Medicine, Johns Hopkins University, Baltimore, MD, USA

(4) Department of Anesthesiology and Critical Care Medicine, School of Medicine, Johns Hopkins University, Baltimore, MD, USA

(5) Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

(6) Department of Oncology Division of Quantitative Sciences, the Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins, Baltimore, MD, USA,

(7) Department of Orthopedic Surgery, School of Medicine, Johns Hopkins University, Baltimore, MD, USA

(*) These authors share senior authorship

Corresponding Author:

Frederick Sieber

Department of Anesthesiology and Critical Care Medicine

Johns Hopkins University School of Medicine

1800 Orleans Street,

Baltimore, MD 21287

Email: fsieber1@jhmi.edu

Main text word count:

Number of data elements:

Abstract:

This study aimed to develop a framework and metrics for characterizing total knee replacement (TKR) stressors and to investigate their link with resilience phenotypes. Key stressor metrics were identified through a sequential elicitation process and categorized as exogenous or endogenous. A subset of stressor metrics was prioritized for analysis based on heterogeneity. To validate the relationship between variation in TKR stressor metrics and previously determined TKR resilience phenotypes: 1) we determined whether pre-TKR phenotypic measures predicted stressor metrics; 2) we determined whether stressor metrics predicted TKR phenotypic trajectories at 1 and 6 months postoperatively. Analyses revealed inconsistent associations between the subset of stressor metrics and pre-TKR resilience phenotype measurements and one-month phenotypic trajectory change. A subset of endogenous metrics (outpatient vs inpatient surgery, intraoperative hypotension, blood loss, vasopressors, and surgical duration) exhibited the expected direction of association with six-month Pittsburgh fatigability scale (PFS) phenotypic trajectory change. This finding indicated that higher stress levels predicted a diminished return of vigor post-surgery. The endogenous metric outpatient vs inpatient surgery was associated with beneficial change in 6-month trajectory of all TKR resiliency phenotype measurements (Short Physical Performance Battery score, PFS, Short Form-36 physical component summary score, knee injury and osteoarthritis outcome score quality of life subscale). This study underscores the importance of stressor variation in resilience research. The developed framework and metrics provide a foundation for future investigations into factors influencing resilience to physical stressors in older adults, particularly highlighting the impact of endogenous factors on post-operative recovery.

Key words: Resilience, phenotype; Surgical procedure, knee replacement; Stressor, metrics

Introduction

Human aging elicits strikingly heterogeneous responses to stressors. Some older adults thrive when encountering adverse life events, challenging clinical procedures, or health insults. Others exhibit lesser capacity for recovery following insults. Understanding what underlies these various responses—and acting to enlarge the proportion who thrive—is among the highest goals of geriatric medicine and gerontology. Biological aging, frailty and reserve are among the gerontological concepts aiming to increase such understanding.

Resilience generally connotes overall potential to emerge with health and functioning intact in the face of stressors (2). Early consensus in gerontology has emphasized that resilience is not a general property but should be referred to more properly as resiliencies—each, specific to a given stressor and indexed by its magnitude as we may be resilient to a minor stressor but not to a more severe one (3). Physical resilience recently has emerged as another like-minded gerontological concept (1). Aligning with this concept, physical resilience is often defined as the ability to adapt and recover well from physical stress (4). However, in the literature on physical resilience in older adults, little has yet been published regarding characterization of the stressor itself. It is important to measure the stressor magnitude and type because they partly determine the post-stressor response, together with an individual's physiologic and other resources for successfully surmounting or rebounding from stressors (henceforth, "resilience capacity") Specifically, the effects of the stressor are hypothesized to be moderated by the resilience capacity (4).

This manuscript presents a conceptual framework for considering the role of stressor metrics and their variation in studies of resiliency. It then applies this framework in a study of total knee replacement (TKR). An elicitation process to identify key stressor metrics and their variation in the TKR context and a taxonomy are proposed. Preliminary study data are used to validate resulting stressor metrics within the proposed framework. This work establishes a foundation for investigations into the factors influencing resilience to physical stressors in older adults.

METHODS

Data source and exemplar stressor: We designed and implemented the *Study of Physical Resilience and Aging (SPRING)* to determine candidate physiological and biological measures by which to identify traits of physical resilience to a clinical stressor / procedure (4). SPRING studies three clinical stressors: total knee replacement (TKR), incident hemodialysis, and bone marrow transplant for hematological cancers (please see (4) for further details). Herein we focus on TKR—the RESilience in TOtal knee REplacement (RESTORE) sub-study, around which a broad range of physiologic measurements were studied to determine resilience capacity. However, the approach to develop a conceptual framework is general (Figure 1 supplemental material; (4)).

The RESTORE study design has been detailed elsewhere (4). In brief, TKR candidates were identified in orthopedic clinics at the Johns Hopkins Bayview Medical Center and University of Maryland Medical Center. Persons consenting participated in detailed assessments of health history, biologic and physiologic parameters, and physical and mental functional status approximately one month before their surgery ("Pre-TKR"), and one, six and twelve months following their surgery. All study protocols were approved by the Institutional Review Boards of the respective institutions.

Conceptual Framework. The SPRING conceptual framework has been previously described (4). For convenience, we reiterate its conceptual diagram here (Figure 1 supplemental material;

(4)) as well as the resilience measurements obtained (Table 1 supplemental material; (4)). In brief, we hypothesize that the health of specific physiologic systems relating to stress response—henceforth, "physiologic resilience capacity"—substantially determines one's capacity to respond resiliently to physical stressors. The major interest is in implications of physiologic resilience for regaining physical and cognitive ("phenotypic") functioning following the stressor and avoiding adverse clinical outcomes. Type and magnitude of the stressor experienced is a critical element operating in this physiologic context: Even with a relatively standardized stressor such as a TKR, example elements which may contribute to greater or lesser procedural stress are the occurrence of medical/surgical complications and the type of anesthesia administered. Persons of equivalent physiologic resilience may exhibit different stressor-related phenotypes or outcomes if the magnitude or nature of their stressor experiences differ. The eventual hypothesis we seek to address is that fitter physiologic resilience blunts the effect of increasing stressor intensity.

In this paper, the initial goal was to develop a framework and propose measures to characterize stressor magnitude and type. To this end we used sequential elicitation processes, followed by assessment of their associations with post-surgery versus pre-surgery phenotypic changes to examine criterion validity of the stressor metrics. The elicitation process that developed the conceptual framework for considering stressor properties consisted of multiple conversations among members of our author group as well as a scientific meeting with the entire SPRING team. We report its findings in the Results section.

Development of Measures

Stressor metrics and resilience phenotypes indicating degree of resilience to stressor: An elicitation process was developed and implemented to identify candidate measures. Guided by the conceptual frameworks addressing stressor properties and the overall goals of SPRING, a questionnaire was created seeking expert input on specific metrics by which to characterize stressor magnitude and type as well as resilience phenotypes most relevant to the particular stressor, aspects of biology or physiology most directly affected by the stressor, and clinical outcomes indicating ability to bounce back from the stressor (Supplement Questionnaire 1). Because the result of this process is an outcome of our work, we report the portions relating to stressor characterization and phenotype selection below.

Potential confounders: In this paper's context, these are variables that may jointly impact resilience phenotype and stressor metrics, and are outside the pathway from stressor to phenotype. Considering our available sample size, the number of confounders that could be controlled while avoiding overfitting was limited. Our team reached consensus on the following set considered particularly apt to affect both stressor metrics and phenotypic outcomes: Age, Charlson Comorbidity Index (CCI)(5), Frail & Prefrail status(6), and history of daily opioid use. Each of these variables except frailty status was self-reported by the participant at pre-TKR. Frailty was assessed by the physical frailty phenotype, which entailed self-report of energy, physical activity, and recent weight loss and three standardized objective assessments. Measured height and weight contributed to the assessment of the phenotypic weight loss criterion. Slowness and weakness criteria were assessed by usual gait speed over a four-meter course and grip strength via a hand-held dynamometer.

Statistical analyses. Baseline participant characteristic distributions were summarized using percentages (categorical variables) or means, medians, standard deviations, interquartile ranges, and ranges (continuous variables). Heterogeneity in characteristics defining stressor magnitude was summarized by interquartile range for continuous variables, with violin plots, and as proportions experiencing a higher-stress condition, for categorical variables. A subset of stressor characteristics was prioritized for further analysis based on extent of heterogeneity and judgment regarding clinical importance.

Four resilience phenotype measures considered particularly salient for total knee replacement, reported below as elicitation results, were employed in analyses. These resilience phenotype measures were: Short Physical Performance Battery score (SPPB), Pittsburgh Fatigability scale (PFS), Short Form-36 physical component summary score (PCS), and knee injury and osteoarthritis outcome score (KOOS) quality of life subscale. To study whether pre-TKR resilience measures predict stressor metrics, associations of each pre-TKR phenotype value (as covariate) with prioritized stressor characteristics (as outcomes) were examined, using regression type appropriate to the characteristic scale (linear for continuous stressor traits, logistic for binary characteristics). A model exploring associations with the potential confounders (mutually adjusted) was also fit. We hypothesized exogenous stressor characteristics to be independent of, and endogenous stressor characteristics to be dependent on, pre-TKR phenotypic measurements. To examine criterion validity of stressor metrics, we assessed their associations with post-surgery versus pre-surgery phenotypic changes. Each phenotype we analyzed was measured continuously: Linear regressions were applied, controlling for potential confounders identified above as well as the baseline level of the respective phenotype. In these analyses, criterion validity is evidenced if higher stressor "magnitudes" are associated with worse resilient responses.

Two analysis fine points bear elaboration. First, regression of phenotype change on phenotype baseline is predisposed to produce negative associations, because of the inclusion of baseline level on both sides of the regression equation ("coupling" of the measure 2–measure 1 difference with measure 1) and regression to the mean. To address this, we utilized a regression methodology that filters out the coupling bias by referencing a person's post-pre surgery change to their change in the absence of surgery (7). Because the change in the absence of surgery is not observed (counterfactual), the analysis requires a sensitivity parameter "k" relating the hypothesized association of baseline and subsequent phenotypes with and without surgery. Here we report analyses with k=1.2 and 1.5—values based on analyses reported in our paper developing the method, using data from a large total knee joint replacement registry (7). As a second fine point, we analyzed changes over two intervals: Six months post-surgery versus baseline, and one month post-surgery versus baseline. The former assesses a resilient potential of greater clinical interest—ultimate improvement, whereas the latter assesses a more proximal impact of the stressor.

RESULTS

Table 1 describes the baseline characteristics and resilience phenotype distributions in the surgical cohort. Approximately two-thirds of subjects were female and self-identified as white, respectively, with a mean age of 69 years. The mean BMI was in the obesity range (32), and 59% were classified as pre-frail. Most subjects self-reported as having good to excellent health. Mean baseline SPPB indicated moderate performance. Mean Koos quality of life scores indicated significant limitations on quality of life caused by knee issues.

Properties of Stressors Identified through the Elicitation Process: Our process to develop a conceptual framework elucidating properties relevant to characterizing stressor magnitude and type in studies of resilience yielded the following major properties:

- Heterogeneity–defined as the variance of continuously scaled stressor characteristics or the uniformity of distribution across categories of nominal or ordinal characteristics. Rationale: Measures without variability do not distinguish stressor experience among individuals.
- Exogeneity–defined as the extent to which a stressor feature conceptually could be randomized to patients irrespective of their intrinsic reserve, or rather inherently reflects their intrinsic reserve (endogeneity). Rationale: Without such a distinction, stressor characteristics may be conflated with determinants, baseline robustness, or physiologic resilience capacity.
- Magnitude–defined as the measured quantity of exposure to a stressor characteristic. Examples include intensity, duration, and accumulation of a particular anesthetic. Rationale: Higher exposures are thought to indicate greater stressor magnitudes.
- Domains—defined as distinguishable elements of the stressor experience, for example, anesthesia versus physical insults of surgery. Rationale: We considered this as a useful organizing principle, and that experienced stress may vary more strongly with heterogeneity in some domains than others.
- Type-defined as qualitatively distinct aspects of the stressor experience within domains. Examples include different anesthesia medications or outpatient versus inpatient procedures. Rationale: Types may exact differential stresses on the patient.
- Timing–defined as timing of stressor metric in relation to the surgical stressor distinguished as pre-surgery, intra-surgery, or directly post-surgery. Rationale: Timing may assist in distinguishing exogeneity versus endogeneity as well as stressor magnitude modifiers, indicators, and outcomes.

Stressor Characteristics Identified through the Elicitation Process: The questionnaire used for initial elicitation of TKR stressor metrics specifically sought to distinguish exogenous versus endogenous elements. Subsequent discussion refined the initial measures proposed and also sought to categorize these by conceptual elements. Table 2 displays the exogenous indices of stressor magnitude in RESTORE. All involved properties of the surgical procedure, such as the type of anesthesia administered or the use/non-use of cement. Only the anesthesia dose was measured as a magnitude rather than a "type."

Table 2 also displays the endogenous stressor characteristics identified in RESTORE. Pre-operative characteristics involved medications to treat pain and anxiety as well as the need for iron supplementation. Intraoperative characteristics reflected physiologic regulation during surgery (e.g. mean arterial pressure) or vasopressor medication administered to manage this, blood loss, and surgical duration. For analytic purposes a vasopressor aliquot was defined as intravenous administration of either 10 mg ephedrine or 100 mcg phenylephrine. Outpatient discharge versus inpatient admission was a post-operative characteristic. Endogenous indices of stress reflect underlying patient vulnerabilities. Conceptually subjects reflecting lack of resilience in physiologic and phenotypic measurements should also demonstrate an increase in the severity of their endogenous stress metrics.

Phenotypes Most Relevant to TKR that were Identified through the Elicitation Process: Mobility-linked features, fatigability as a major potential impact of surgery, and overall quality of life were deemed as particularly salient phenotypes. Following extensive discussion, four measures were prioritized: the Short Physical Performance Battery score (SPPB)(8), the Pittsburgh fatigability scale (PFS)(9), the Short Form-36 (SF–36) physical component summary score(10), and the knee injury and osteoarthritis outcome score (KOOS) overall quality of life subscale(11).

Except for mobility, which is a particularly fundamental outcome for TKR and for older adults, these measures sought to identify holistic potential impacts of the TKR stressor, rather than ones tied closely to the knee replacement outcome (Laskow). Importantly, they are intended as "trajectories" - not only baseline function but pre and post information spanning baseline through some period of follow up, where changes are thought to be indicative of resilience.

Analyses of stressor characteristic heterogeneity: There was little variability in exogenous stressor metrics except for intraoperative anesthesia dose (Figure 1A; Table 2).

Among the preoperative endogenous stressor characteristics measured, approximately 1/4 of subjects were chronically managing their pain with gabapentin and 18% were taking anxiolytics (Table 2). Intraoperatively, the median number of minutes during the procedure in which the patient was hypoxic, tachycardic, hypotensive, or hypertensive was minimal, and the greatest variability occurred with minutes of hypotension (Table 2). There was also a wide range of variability in vasopressor aliquots administered, surgical duration, and recorded blood loss (Figure 1B-D). In those patients undergoing inpatient procedures (n=71) there was large variability in postoperative opioid consumption.

Stressor characteristics were prioritized for further analysis as outcomes (of baseline characteristics) and predictors (of phenotypic change) based on offering considerable heterogeneity and, among such measures, their clinical importance. Only anesthesia dose among exogenous measures met these criteria: This was analyzed as average intraoperative infusion rate (mcg/kg/min) of propofol among the 91 individuals who received spinal anesthesia. Endogenous measures carried forward for further analysis were hypotension (dichotomized as ≥2 versus <2 minutes below 60 torr of mean arterial pressure; n=47 per group), estimated ml blood loss (log-transformed), vasopressor aliquots (count plus one log-transformed), surgical duration (in hours), and outpatient discharge versus inpatient admission.

Pre-TKR Phenotype measurements (Resilience Capacity) as predictors of stressor metrics (Table 3):

Few associations between pre-stressor resilience phenotypic measures and stressor metrics were identified. The strongest associations were observed between lower levels of baseline fatigue (measured by reverse-coded PFS score) and a higher chance of having outpatient vs. inpatient surgery and, contrary to our hypothesis, administration of higher doses of anesthesia medication (propofol) during surgery (Table 3). Weaker associations were observed between lower baseline quality of life (KOOS score) and longer surgical duration, and between higher baseline physical health scores (PCS) and a lower chance of needing to stay overnight (inpatient) after surgery (Table 3).

In the series of models characterizing stressor variability in terms of potentially confounding variables, only an isolated association of frail versus robust phenotype with larger intraoperative vasopressor requirements was identified (Supplementary Table 2).

Stressor measurements as predictors of phenotypic resilience trajectories (Table 4):

Only isolated, inconsistent associations between stressor metrics and one-month phenotypic trajectory change were identified, after adjusting for potentially confounding variables (data not shown). A number of consistent adjusted associations were observed between stressor metrics and six-month phenotypic trajectory change. All endogenous measurements analyzed exhibited the expected direction of association with the reverse-coded PFS change from pre-TKR to 6 months, such that values indicating higher stress levels predicted a diminished return of vigor post-surgery (Table 4). Blood loss was most strongly implicated with the reverse-coded PFS change from pre-TKR to 6 months. Outpatient vs inpatient procedures were consistently associated with more beneficial change from baseline to 6 months of all resiliency phenotype measurements. Impact on the SPPB score recovery was most strongly evidenced. Other individual strong associations were observed, but these occurred with less consistency across phenotypic trajectories or stressor metrics. Increased vasopressor aliquots were negatively associated with each of the self-reported phenotypes but only negligibly associated with the SPPB score. The SPPB score recovery was negatively associated with >2 minute hypotension duration and with longer surgical duration, in addition to outpatient vs inpatient procedure.

To determine the association between potential confounders and phenotypic trajectories, pre-TKR resilience phenotype measure were controlled for in mutually-adjusted analyses. The Charlson comorbidity index (CCI) was associated with poorer one-month rebound and six-month rebound in both the reverse-coded PFS score and the KOOS quality of life score. CCI also was associated with poorer Physical PCS rebound. Less consistent associations were observed with older age (worse six-month rebound in PFS and Physical PCS), history of daily opioid use (worse one-month rebound in PFS), and frail vs robust status (worse six-month rebound in SPPB score) (Supplementary Table 3).

DISCUSSION

A conceptual framework was presented for determining the role of variation in stressor metrics in studies of resiliency using TKR as an example. A taxonomy approach highlighting endogenous and exogenous mechanisms was used to identify and characterize the stressor metrics of TKR. Using the resilience phenotypes of TKR, it was investigated: 1) whether pre-

TKR phenotypic measures predicted stressor metrics; 2) whether stressor metrics predicted phenotypic trajectories at various follow-up time points. Only inconsistent associations between pre-TKR phenotype measures and stressor metrics or stressor metrics and one-month change in resilience phenotypes were observed. More consistent associations were identified between stressor metrics and 6-month change in resilience phenotypes.

We consider the distinction between endogenous and exogenous stressor characteristics to be important. Exogenous metrics, factors outside the physiological system we are studying, may be more amenable to interventions aimed at improving resilience and outcomes. However, in the RESTORE study, most exogenous stressors were similar because surgery and treatment plans were standardized. The propofol infusion rate used during surgery (an exogenous factor) varied the most, as propofol dosing is not subject to strict guidelines. Analogously, exogenous variation could be substantial across institutions or medical practices where guidelines may differ. Commensurately, multi-site designs would be valuable in future research on physical resilience.

Endogenous metrics, rooted in a patient's individual characteristics, are inherently more complex and varied than exogenous factors. This pattern was observed in total knee replacement (TKR). Intrinsic factors, including genetic predisposition, psychological traits, and physiological responses, can significantly influence a patient's resilience to stressors. While exogenous factors can be controlled or manipulated, endogenous metrics are often more difficult to modify directly. One potential strategy for enhancing resilience in the context of endogenous factors is to avoid extremes. This might involve identifying individual vulnerabilities. Assessing a patient's genetic predispositions, psychological traits, or physiological responses allows healthcare providers to identify areas where the patient may be particularly susceptible to stress. Based on this assessment, personalized interventions can be developed to mitigate the impact of these vulnerabilities. Our study identified fatigability as an outcome that may be particularly susceptible to stressor effects: Interventions to enhance energy pre-surgically may offer a promising direction in the TKR setting. By understanding and addressing endogenous factors, healthcare providers can develop more effective strategies for improving resilience and supporting patients in their recovery from stressors.

This study investigated the connection between pre-TKR resilience phenotype measurements and stressor metrics. Overall, few clear links between pre-TKR resilience phenotypes and stressor response emerged. The study acknowledges that the measures of resilience we employed might not fully reflect underlying health problems, other vulnerabilities, or resilience capacity which may account for this discrepancy. Alternatively, factors we identified as endogenous may be less dependent on personal characteristics (and more dependent on surgical practice) than we hypothesized.

The long term recovery trajectory of the resilience phenotypes is a function of demographics, stress response, resilience capacity, and other additional features such as environmental factors (e.g., social support). In our validation analysis, then, we hypothesized that "outcomes" of stressor variation should operate on a shorter time scale and could be transient, and that stressor impacts on longer term recovery trajectories are blunted by the many factors that may modify these. The data did not sustain these hypotheses. It is possible that the baseline-1 month trajectory change was not a short enough time frame, or its measurement was too variable, to observe consistent associations with stressor metrics. It also is possible that our prioritized stressor metrics did not include the strongest short-term actors. On the other hand, multiple consistent signals of stressor variation impacts on six-month recovery trajectories were observed.

Associations with the outpatient vs inpatient stressor metric were in the expected direction for all phenotypes considered. Higher levels of each stressor we considered were associated with a diminished six-month return of vigor post-surgery. Higher vasopressor exposure was associated with considerably diminished recovery in all three of the self-reported phenotypes we studied. Associations such as these highlight that effects of pre-surgical determinants and capacity on post-surgical recovery may be mediated through endogenous stressor variation—an important analytic consideration for future studies of resilience. Selection biases and confounding, alternatively, may be at play: Further study is warranted to illuminate causality. In summary, our study evidences that the stressor-characterizing indicators we studied are associated with longer term TKR recovery, hence capture clinically relevant variation in stressor magnitude or type.

The phenotypic measurements we studied are both self-reported (PFS-rev, physical PCS, KOOS-QOL) and measure function objectively (SPPB). Our pre-TKR phenotypic measures indicate pre-reserve, suggesting a level of functional capacity prior to surgery. However, the relationship between self-reported pre-reserve and knee-specific outcomes remains unclear. Furthermore, trajectories of the phenotypes we studied and outcomes specific to the knee area may be concordant or discordant. It is possible that individuals with a lower degree of concordance between phenotypic trajectories and knee specific outcomes may report a lower level of pre-reserve, even though their actual functional capacity may be comparable to those with higher concordance. Further research is needed to investigate the interaction between patient-reported measurements of pre-reserve and concordance/discordance between phenotypic trajectories and knee specific outcomes.

Study strengths and limitations:

Our study is among the first proactively designed to study older adults' resilience to physical stressors. Systematic deliberation was applied to develop a taxonomy characterizing stressor characteristics. Data in RESTORE were collected according to rigorously standardized, research-directed protocols and at the same time are fully representative of a real-world clinical setting. An innovative statistical methodology designed to mitigate mathematical artifacts inherent to analyses regressing phenotypic change on phenotypic baseline was employed.

Among limitations, the study suffers from limited sample size. This limited power to detect associations. Our interpretation of findings therefore emphasized consistent patterns, commensurate with an early-stage study. Future validation will be needed. Among the outcome types we considered, associations with six-month phenotypic change were solidly evidenced, accounting for multiple comparisons: The number of findings significant at a p=0.10 level, for example, was nearly three times that expected by chance. Our sample size also limited the extent of control for potential confounders that could be achieved. Although there was strong phenotypic heterogeneity, questions remain concerning success in recruiting the least resilient individuals. In determining the stressor metrics there was necessarily some ambiguity in determining endogenous/exogenous stressor elements. However, this problem was limited by focusing on the ability to randomize exogenous elements as the important selection criteria. Variability in exogenous stressor elements could result from confounding by indication. For instance, anesthesiologists may tune anesthetic dosing to perceived patient vulnerability. The analysis reported herein, finally, assesses stressor measure via criterion validity: We defer analysis of the primary study hypothesis that higher resilience capacity buffers stressor impacts to subsequent work. Our study's limitations are inherent to early-stage research and primarily render our findings incomplete. We believe that the conceptual work accomplished and our findings relating to six-month recovery provide necessary foundation for future research directions and study designs, and potentially important clinical insights.

In summary, our study provides a conceptually rigorous taxonomy for stressor characterization in the context of resilience studies. The SPRING study was designed to determine candidate physiological and biological measures by which to identify traits of physical resilience to three clinical stressors: total knee replacement (TKR), incident hemodialysis, and bone marrow transplant for hematological cancers. An elicitation process to identify candidate stressor and phenotypic measures was developed and implemented. TKR study data were used to validate measures of stressor magnitude. This approach may be generalizable to other stressors and studies of physical resilience.

LEGENDS:

Figure 1 : A) Density curve and box plot of the exogenous stressor characteristic of average propofol dose administered intraoperatively (mcg/kg/min) in those subjects undergoing spinal anesthesia. Propofol dose had greatest variability of the exogenous stressor characteristics measured. B,C,D) Logarithmically transformed density curves and box plots of endogenous stressor characteristics with greatest intraoperative variability. B) Log blood loss; C) Log total vasopressor aliquots; D) Log duration of surgery in hours. Box plot rectangles denote first and third quartiles and central dots the median.

Funding:

This work was supported by grant UH3AG056933 from the National Institute of Aging

Acknowledgements:

The authors would like to thank Meredith Dobrosieski who recruited, consented, and obtained all outpatient measurements and Armen Davani who obtained all in-hospital measurements. In addition, we thank the orthopedic surgeons at Johns Hopkins Bayview Medical Center (Harpal Khanuja, Robert Sterling, Vishal Hegde) for their continued support of this study.

Conflicts of Interest: None

Preprint status: This work has not been peer reviewed.

REFERENCES:

1. Resnick B, Galik E, Dorsey S, Scheve A, Gutkin S. Reliability and validity testing of the physical resilience measure. Gerontologist. 2011;51:643–652.

2. Whitson HE, Duan-Porter W, Schmader KE, Morey MC, Cohen HJ, Colón-Emeric CS. Physical Resilience in Older Adults: Systematic Review and Development of an Emerging Construct. J Gerontol A Biol Sci Med Sci. 2016;71:489–495.

3. Hadley EC, Kuchel GA, Newman AB, Workshop Speakers and Participants. Report: NIA Workshop on Measures of Physiologic Resiliencies in Human Aging. J Gerontol A Biol Sci Med Sci. 2017;72:980–990.

4. Walston J, Varadhan R, Xue Q, et al. A Study of Physical Resilience and Aging (SPRING): Conceptual framework, rationale, and study design. J Am Geriatr Soc. 2023;71:2393–2405.

5. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis. 1987;40:373–383.

6. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci. 2001;56:146.

7. Varadhan R, Zhu J, Bandeen-Roche K. Identifying predictors of resilience to stressors in single-arm studies of pre-post change. Biostatistics. 2023.

8. Guralnik JM, Simonsick EM, Ferrucci L, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol. 1994;49:85.

9. Glynn NW, Santanasto AJ, Simonsick EM, et al. The Pittsburgh Fatigability scale for older adults: development and validation. J Am Geriatr Soc. 2015;63:130–135.

10. Laucis NC, Hays RD, Bhattacharyya T. Scoring the SF-36 in Orthopaedics: A Brief Guide. J Bone Joint Surg Am. 2015;97:1628–1634.

11. Roos EM, Toksvig-Larsen S. Knee injury and Osteoarthritis Outcome Score (KOOS) - validation and comparison to the WOMAC in total knee replacement. Health Qual Life Outcomes. 2003;1:17–17.

Figure 1 a.

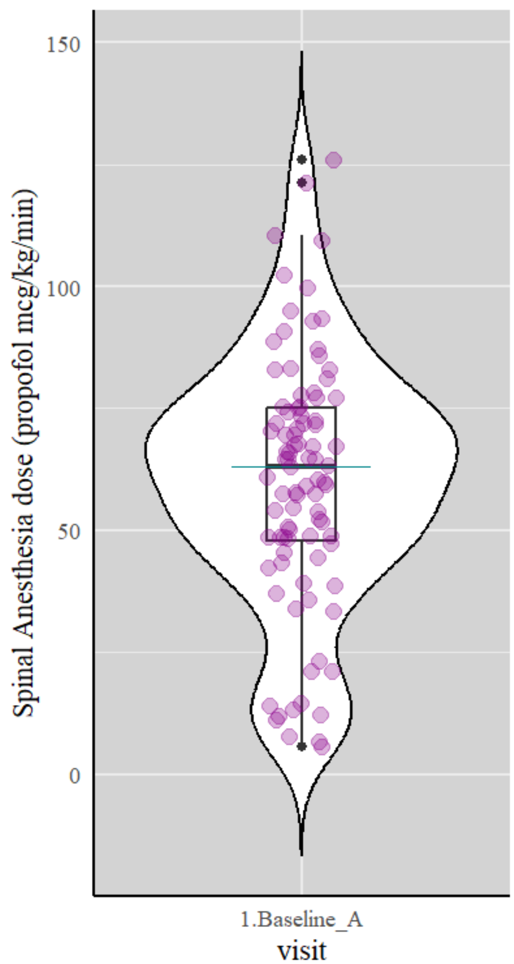


Figure 1 b.

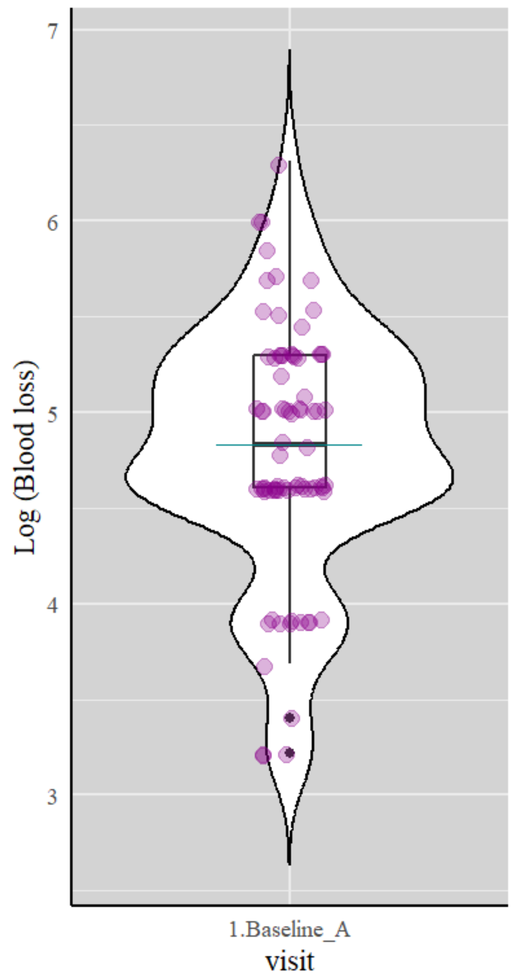


Figure 1 c.

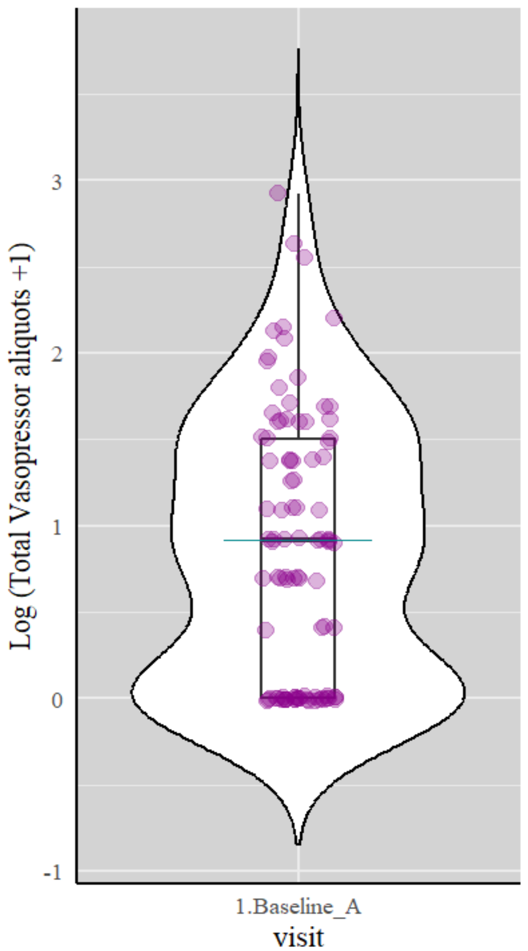


Figure 1 d.

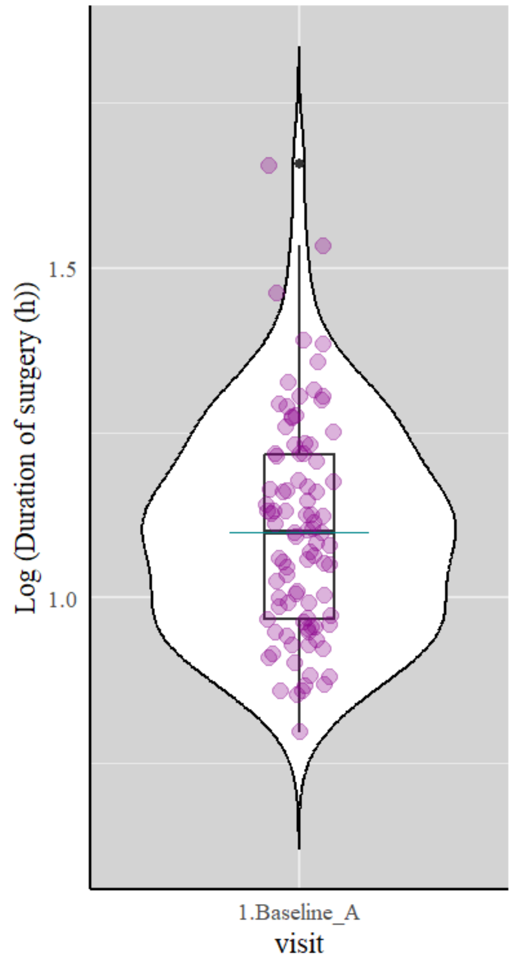


Table 1. Pre-TKR Characteristics of Study Sample (N=94).

Age	
Mean (SD)	69.1 (6.8)
Median	68.5
Range	(59.0, 85.0)
Sex, n (%)	
Female	63 (67)
Male	31 (33)
Race, n (%)	
White	58 (62)
Black	33 (35)
Other	3 (3)
Education in year	
Mean (SD)	14.5 (2.8)
Median	14.0
Range	(7.0, 20.0)
Body Mass Index	
Mean (SD)	32.1 (5.6)
Median	31.5
Range	(20.0, 51.0)
Marital status, n (%)	
Never Married	8 (9)
Widowed	22 (23)
Separated/Divorced	12 (13)
Married/ Living with someone	52 (55)
History of daily opioid use, n (%)	
yes	22 (23)
no	72 (77)
Charlson Comorbidity Index, n (%)	
None	54 (57)

Mild	25 (27)
Moderate/Severe	15 (17)
Physical Frailty, n (%)	
Non-frail	24 (25.5)
Prefrail	55 (58.5)
Frail	15 (16)
Self-reported health, n (%)	
Excellent/Very good	38 (40)
Good	37 (39)
Fair/Poor	19 (20)
***Gait speed (m/s)**	
Mean (SD)	0.8 (0.2)
Median	0.8
Range	(0.2, 1.5)
Total Score of the Short Physical Performance Battery (SPPB)	
Mean (STD)	8.3 (2.3)
Median (IQR)	8.0 (7.0, 10.0)
Range	(3, 12)
Physical fatigability Score (PFS) "Reversed"	
Mean (STD)	31.1 (10.8)
Median (IQR)	31.5 (23.0, 39.0)
Range	(10, 50)
SF-36 Physical Component Summary (PCS)	
Mean (STD)	31.3 (8.5)
Median (IQR)	29.6 (25.9, 38.5)
Range	(13, 48)

KOOS Quality of Life Score (QOL)	
Mean (STD)	27.3 (18.5)
Median (IQR)	25.0 (12.5, 37.5)
Range	(0, 75)

Note: There were no missing values for frailty or self-rated health (SRH). However, gait speed data was missing for two study IDs.

Table 2: Exogenous and Endogenous Indices of Stressor Magnitude of Total Knee Replacement.

Stressor Type	Stressor timing	Characteristic	Variability
Exogenous	Intra- operative	Tranexamic acid administration vs. not	93 **vs.** 1
		Spinal Anesthesia **vs.** General Anesthesia	91 **vs.** 3
		Average infusion rate of propofol intraoperatively during Spinal Anesthesia [a]	60.09 (± 26.01) [5.60, 125.86]
		Total vs Unicompartment Knee Replacement	94 **vs.** 0
		Use of tourniquet **vs**. no tourniquet	94 **vs.** 0
		Use of cement **vs.** no cement	81 **vs.** 13
Endogenous	Pre- operative	Receiving **vs.** not receiving gabapentin	24 **vs**. 70
		Receiving **vs.** not receiving anxiolytics	17 **vs.** 77
		Requirement for iron therapy **vs.** not	9 **vs.** 85
	Intra- operative	Duration of hypotension (MAP < 60 torr) [b]	1.00 (0.00 – 11.00) [0.00, 93.00]
		Duration of hypertension (SBP > 160 torr) [b]	0.00 (0.00 – 5.00) [0.00, 36.00]
		Duration of pulse rate < 100 BPM [b]	0.00 (0.00 – 0.00) [0.00, 61.00]
		Duration of hypoxia (O2 saturation < 90%) [b]	0.00 (0.00 – 0.00) [0.00, 11.00]
		Total Vasopressor aliquots (ephedrine 10 mg or phenylephrine 100 mcg) [c]	1.50 (0.00 – 3.50) [0.00, 17.50]
		Blood loss [d]	122.50 (100.00 – 200.00) [0.00, 550.00]
		Duration of surgery [e]	3.1 (±0.5) [2.22, 5.25]
	Post-operative	Outpatient procedure **vs.** inpatient procedure	23 **vs.** 71
		Postoperative day 1 opioid consumption for inpatients in morphine equivalents.	12.50 (7.50 – 22.50) [0.00, 132.50]

		Discharge disposition: home/self-care **vs.** home/health care services **vs.** skilled nursing facility	78 **vs.** 15 **vs.** 1
		Medical/surgical complications: none **vs.** minor **vs.** major	73 **vs.** 14 **vs.** 7

Note: For binary characteristics – variability is given as number with versus without characteristic. For continuously measured characteristics – values are given as mean (± standard deviation) [minimum, maximum] or median (IQR) [minimum, maximum].

Note: MAP : Mean Arterial Pressure. SBP : Systolic Blood Pressure. BPM : Beats Per Minute. O2 : Oxygen. Mcg/kg/min :micrograms per kilogram per minute. Stressor characteristic measurement units : [a] mcg/kg/min [b] minutes [c] number of doses [d] milliliters [e] hours.

Table 3: Associations of Pre-TKR Resilience Phenotype Measures with Stressor Metrics (Separate models, n=94).

		Pre-TKR Resilience Phenotype Measures			
Stressor Type	Stressor Metric	SPPB	PFS (Reversed)	SF36 PCS	KOOS QOL
Endogenous	Hypotension>2 min[a]	1.02 (0.85, 1.22)	0.98 (0.95, 1.03)	0.99 (0.95, 1.05)	0.99 (0.98, 1.02)
	Blood loss	0.01 (-0.06, 0.07)	0.01 (-0.01, 0.02)	-0.00 (-0.02, 0.01)	-0.00 (-0.01, 0.01)
	Vasopressor aliquots	0.01 (-0.06, 0.08)	0.00 (-0.01, 0.02)	0.01 (-0.01, 0.03)	0.01 (-0.00, 0.02)
	Surgical duration	-0.00 (-0.02, 0.01)	0.00 (-0.00, 0.00)	-0.00 (-0.01, 0.00)	-0.001 (-0.01, 0.00) *
	Inpatient admit [a]	0.87 (0.70, 1.08)	0.94 (0.89, 0.98) **	0.95 (0.90, 1.01) *	0.98 (0.95, 1.01)
Exogenous	Anesthesia dose	1.55 (-0.83, 3.94)	0.63 (0.14, 1.12) **	0.14 (-0.49, 0.78)	-0.13 (-0.42, 0.16)

Note: Analyses are based on unadjusted models, each performed separately for phenotypic measures such as SPPB, PFS (reversed), SF-36 PCS, and KOOS Quality of Life. Continuous stressor characteristics analyzed linearly underwent specific transformations: Blood loss and surgical duration were subjected to logarithmic transformations; vasopressor aliquots were logarithmically adjusted by adding 1; and average propofol infusion rate during spinal anesthesia measurements were conducted in mcg/kg/min units without modification. [a] Results are given as OR (95% CI). $**p<0.05$. $*p<0.10$.

Note: Total score of the Short Physical Performance Battery (SPPB), Physical Fatigability Score (PFS) "Reversed", SF-36 Physical Component Summary (PCS), and the Knee Injury and Osteoarthritis Outcome Score (KOOS) - Quality of Life subscale.

Table 4. Associations of stressor characteristics with six-month change in resilience phenotypes in the Study of Physical Resilience in Aging Pilot.

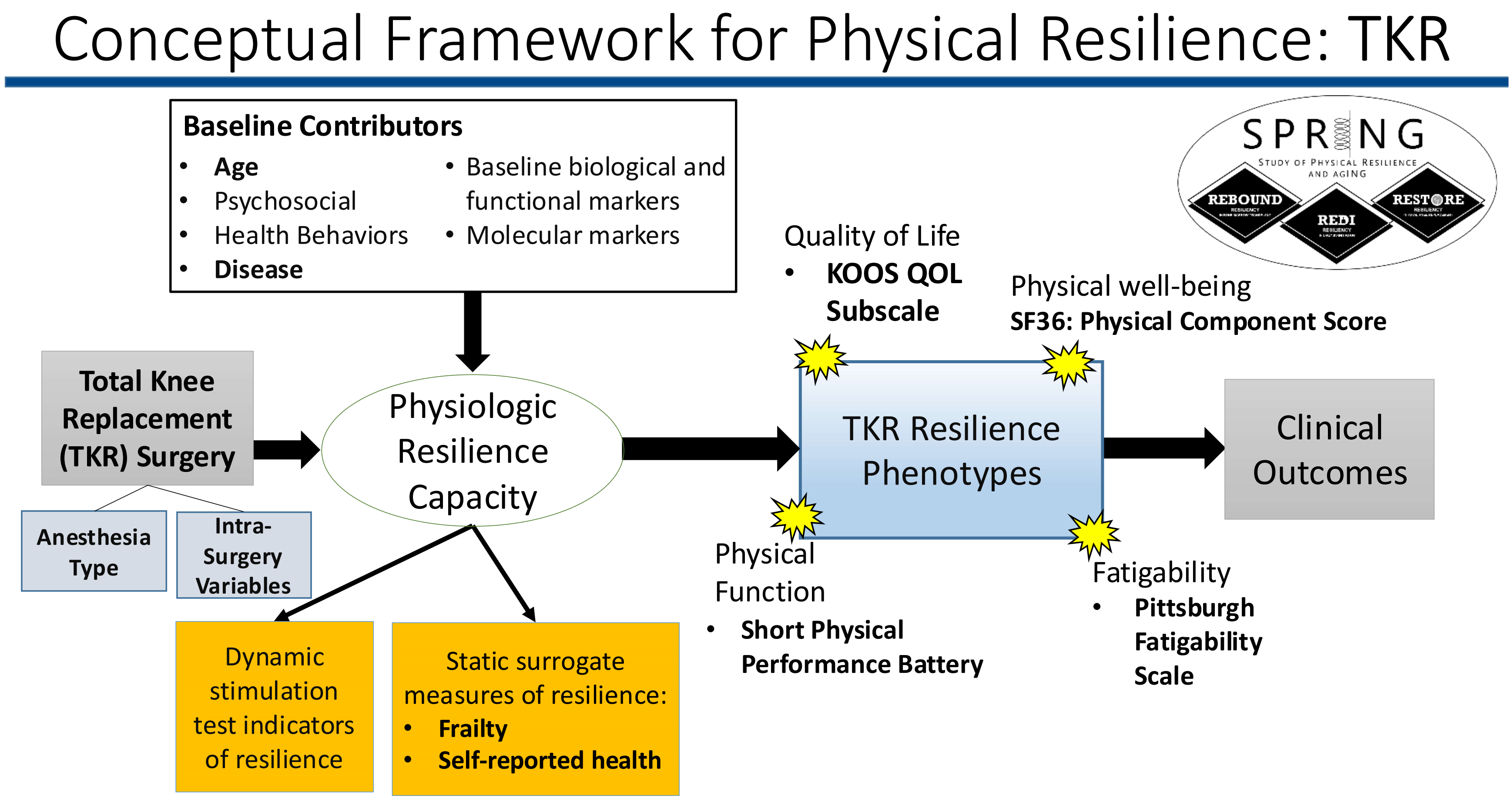
Stressor Type	Stressor Characteristic	k	PFS (Reversed)	SPPB	Physical PCS	KOOS-QOL
Endogenous	Hypotension > 2 min	1.2	-2.74 (-7.10, 1.02)	-0.64 (-1.44, 0.14)*	1.06 (-2.79, 4.87)	-4.24 (-14.85, 6.06)
		1.5	-2.35 (-6.48, 1.20)	-0.60 (-1.33, 0.13)*	0.90 (-2.65, 4.57)	-4.19 (-14.35, 5.69)
	Blood loss	1.2	-2.81 (-6.23, -0.09)**	0.08 (-0.61, 0.79)	-0.78 (-3.55, 1.87)	2.73 (-4.90, 11.51)
		1.5	-2.44 (-5.66, 0.34)*	0.07 (-0.54, 0.72)	-0.86 (-3.55, 1.74)	2.43 (-4.70, 10.96)
	Vasopressor aliquots	1.2	-1.80 (-4.63, 0.65)	0.04 (-0.50, 0.58)	-2.75 (-5.33, -0.19)**	-6.30 (-13.40,1.35)*
		1.5	-1.69 (-4.26, 0.60)	0.10 (-0.40, 0.59)	-2.47 (-5.09, 0.07)*	-5.93 (-12.75,1.25)*
	Surgical duration	1.2	-2.08 (-5.33, 1.85)	0.68 (-0.03, 1.78)*	0.39 (-2.77, 3.01)	4.59 (-4.72, 14.59)
		1.5	-1.81 (-4.86, 1.75)	0.55 (-0.16, 1.64)*	-0.01 (-2.92, 2.53)	4.00 (-4.91, 14.05)
	Outpatient admit	1.2	1.68 (-2.91, 6.35)	0.90 (0.05, 1.86)**	2.67 (-1.12, 6.59)	5.49 (-7.28, 17.19)
		1.5	2.61 (-1.74, 6.57)	0.90 (0.12, 1.73)**	3.16 (-0.64, 6.64)*	6.12 (-6.27, 17.19)
Exogenous	Anesthesia dose	1.2	-0.06 (-0.17, 0.03)	<0.01 (-0.02, 0.02)	-0.01 (-0.07, 0.05)	0.07 (-0.20, 0.31)
		1.5	-0.04 (-0.16, 0.05)	<0.01 (-0.02, 0.02)	-0.01 (-0.08, 0.05)	0.06 (-0.19, 0.30)

Note: n = 89 for all outcomes except SPPB and predictors except Anesthesia dose and Blood loss. For SPPB, n = 79 for all predictors but Anesthesia dose and Blood loss. For Anesthesia dose, n=86 for outcomes other than SPPB and = 76 for SPPB. For Blood loss, n = 77 for outcomes other than SPPB and = 71 for SPPB.

Note: Analysis was linear regression correcting for regression to the mean of change on baseline using the method of Varadhan et al. 2023. k is a sensitivity parameter for this method, varied here over two plausible values.

Note: Analyses are adjusted for age (in years), Charlson comorbidity index score, frail and prefrail status (each versus robust), and pre-surgery opioid use. **$p < 0.05$ & 95% confidence intervals do not include 0. *Indicates cases, where the 95% confidence intervals include 0, meaning they are not significant at the $p < 0.05$ level, but the overlap beyond 0 is less than 10% of the total width.

Supplemental Figure 1.



Supplemental Figure 1: The conceptual organizing framework that guides measurement collection in the total knee replacement (RESTORE) portion of the SPRING study. In this framework, the clinical stressor, TKR, elicits a resilience response measured by resilience phenotypes (blue box). Resilience phenotypes characterize both the initial response to the stressor and the recovery or decline thereafter in key markers of physical and cognitive function and health. These therefore are “trajectories” of key measures, beginning shortly before TKR and continuing afterwards up until a time frame at which we may expect a resilient person to recover their pre stressor status- operationalized in this study as one year. We hypothesize that physiologic resilience capacity is influenced by static pre- TKR measures (white box) and can be elicited by measures stressing the functionality of specific dynamical systems that are measured before the stressor (left orange Box). We also hypothesize that physical frailty and self-reported health can function as surrogate measures of resilience capacity (right orange box). Both the resilience capacity and phenotypes potentially contribute to clinical outcomes (gray box). Adapted from J Am Geriatr Soc. 2023;71:2393-2405.

Supplementary Table 1: Measurements	
Resiliency Measurement Categories	**Total Knee Replacement**
Determinants	Age, demographics, SES, vitals, anthropometrics, psychosocial measures, health behaviors, medical history, cell senescence
Stressor Magnitude	anesthesia type & dose, time in surgery, blood loss, fluids administered during surgery, transfusion and/or pressor use, surgical procedure, perioperative drugs administered, tourniquet use, outpatient vs inpatient surgery, discharge disposition, NSQIP complications
Surrogates*	Phenotypic frailty, self-reported health, Karnofsky performance scale, nutrition (albumin, total cholesterol, vitamin D)
Dynamic Stimulation Measures*	ACTH stimulation, diurnal salivary cortisol profile, oral glucose tolerance test, Holter monitoring, observed fatigability, dynamic ex-vivo response of immune cells, orthostatic blood pressure
Phenotypic Indicators*	Physical function (SPPB, mobility function), accelerometry, cognitive function (MOCA, NIH Toolbox cognition measures), disability, fatigability scale, spirometry, SF-36 physical component, KOOS-QOL
Physiologic Indicators	Physiologic markers*: clotting factors, hormones, inflammatory markers, immune cell phenotypes, overnight catecholamines, Angiotensin System auto-antibodies, metabolomic markers, reactivation of CMV, oxygen saturation, end tidal CO2, and body temperature. Dynamic assessments#: Cerebral autoregulation, heart rate and blood pressure variability
Outcomes	Hospital readmit, falls, mortality, length of stay until discharge, delirium, length of rehab, ability to ambulate unaided 30/60 days later, KOOS pain

Supplementary Questionnaire 1

Questions used to help characterize Restore stressor magnitude:

1) What are characteristics by which we might characterize the magnitude of TKR stressor in RESTORE? Please think of "exogenous" aspects of TKR which, in a twisted world, might be "randomized" to patients, rather than "endogenous" characteristics of the patient's own vulnerabilities.

2) Which resilience phenotypes are most relevant to the total knee replacement clinical stressor?

3) Which aspects of biology / physiology are most directly affected by TKR?

4) What are primary clinical outcomes indicating ability to bounce back from TKR?

Supplementary Table 2: Associations of Demographic/Health Characteristics with Stressor Metrics (n=94).

Stressor Type	Stressor Metric	Potential Confounders: Age	Charlson Comorbidity Index	Frail vs. Robust	History of Daily Opioid Use vs. None	*p*-value
Endogenous	Hypotension>2 min[a]	0.98 (0.92, 1.05)	1.13 (0.82,1.55)	0.31 (0.07, 1.35)	1.91 (0.67, 5.41)	0.474
	Blood loss	0.00 (-0.02, 0.02)	-0.03 (-0.14, 0.09)	-0.01 (-0.57, 0.56)	0.03 (-0.38, 0.43)	0.975
	Vasopressor aliquots	0.00 (-0.02, 0.02)	-0.08 (-0.20, 0.03)	0.63 (0.12, 1.15)*	-0.25 (-0.62, 0.12)	0.180
	Surgical duration	-0.00 (-0.01, 0.00)	-0.00 (-0.03, 0.02)	-0.01 (-0.12, 0.10)	-0.05 (-0.13, 0.03)	0.569
	Inpatient admit [a]	1.06 (0.98, 1.14)	1.31 (0.84, 2.05)	3.83 (0.39, 37.70)	2.36 (0.58, 9.50)	0.077
Exogenous	Anesthesia dose	-0.28 (-1.13, 0.56)	-1.07 (-5.28, 3.13)	-2.42 (-20.79, 15.95)	-6.06 (-19.42, 7.31)	0.861

Note: Analyses are adjusted for age in years, Charlson comorbidity index score (CCI), frail and prefrail status (each; versus robust); and use of opioids pre-surgery. Continuous stressor characteristics analyzed linearly underwent specific transformations: Blood loss and surgical duration were subjected to logarithmic transformations; vasopressor aliquots were logarithmically adjusted by adding 1; and average propofol infusion rate during spinal anesthesia measurements were conducted in mcg/kg/min units without modification.

Note: Chi-square tests were applied for Hypotension > 2 min and Inpatient vs. Outpatient, while F-tests were used for Blood loss, Vasopressor aliquots, Surgical duration, and Anesthesia dose to evaluate the null hypothesis of no association between outcomes and covariates. The table displays the likelihood ratios or F-values alongside their respective p-values (Pr > ChiSq or Pr > F).

Note: Demographics-only model includes the pre-determined confounders of age, Charlson Comorbidity Index, frail status, and opioid use pre-surgery. [a] Results are given as OR (95% CI). *p<0.05. **p< 0.10.

Note: Total score of the Short Physical Performance Battery (SPPB), Physical Fatigability Score (PFS) "Reversed", SF-36 Physical Component Summary (PCS), and the Knee Injury and Osteoarthritis Outcome Score (KOOS) - Quality of Life subscale.

Supplementary Table 3: Associations of demographic/health characteristics with six-month change in resilience phenotypes in the study of physical resilience in aging pilot.

		Resilience Phenotype			
Confounder	k	PFS-rev	SPPB	Physical PCS	KOOS-QOL
Age	1.2	-0.32 (-0.61, -0.04)*	<0.01 (-0.05, 0.08)	-0.24 (-0.53, 0.05)**	-0.32 (-1.16, 0.49)
	1.5	-0.32 (-0.58, -0.06)*	-0.02 (-0.08, 0.05)	-0.21 (-0.47, 0.07)	-0.26 (-1.10, 0.52)
Charlson Comorbidity Index	1.2	-1.20 (-2.47, -0.02)*	-0.02 (-0.43, 0.28)	-0.66 (-1.80, 0.42)	-3.82 (-8.79, -0.42)*
	1.5	-1.36 (-2.59, -0.17)*	-0.03 (-0.41, 0.24)	-0.99 (-2.18, 0.04)**	-3.75 (-8.58, -0.50)*
Prefrail vs. Robust	1.2	1.12 (-3.46, 5.73)	0.46 (-0.40, 1.45)	1.13 (-2.63, 5.65)	-6.31 (-18.7, 6.33)
	1.5	0.43 (-3.89, 5.12)	0.11 (-0.65, 1.17)	0.41 (-3.16, 4.67)	-6.34 (-18.2, 5.83)
Frail vs. Robust	1.2	2.67 (-4.64, 10.7)	2.03 (0.62, 3.63)[a]*	-1.85 (-6.83, 3.87)	7.52 (-11.5, 26.6)
	1.5	0.62 (-5.88, 8.35)	1.32 (0.05, 3.16)*	-2.57 (-7.42, 2.74)	6.71 (-11.5, 25.4)
History of Daily Opioid Use vs. None	1.2	-2.65 (-7.45, 3.39)	0.05 (-1.23, 1.02)	-2.22 (-6.71, 2.84)	4.88 (-8.09, 17.7)
	1.5	-3.74 (-8.39, 2.04)	-0.17 (-1.33, 0.75)	-3.42 (-7.95, 1.61)	4.60 (-7.68, 17.2)

Notes: a) Percentile bootstrap interval: BCA interval did not converge.

Note: Total score of the Short Physical Performance Battery (SPPB), Physical Fatigability Score (PFS) "Reversed", SF-36 Physical Component Summary (PCS), and the Knee Injury and Osteoarthritis Outcome Score (KOOS) - Quality of Life subscale. Covariates: age, Charlson Comorbidity Index, frail status, and opioid use pre-surgery. *p < 0.05 & 95% confidence intervals do not include 0. **Indicates cases, where the 95% confidence intervals include 0, meaning they are not significant at the p < 0.05 level, but the overlap beyond 0 is less than 10% of the total width.