

Applying Allostatic Load to Perinatal Outcomes Research

by

Yang Li

**A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Nursing)
in the University of Michigan
2018**

Doctoral Committee:

**Professor Julia S. Seng, Chair
Professor Vanessa K. Dalton
Associate Professor Shawna J. Lee
Assistant Professor Marie-Anne S. Rosemberg
Associate Research Scientist James J. Yang**

Yang Li

lyx@umich.edu

ORCID iD: 0000-0001-8901-3454

© Yang Li, 2018

ACKNOWLEDGEMENTS

First of all, I would like to express my deepest gratitude to my mentor, Dr. Julia Seng. I am tremendously fortunate to have Dr. Seng as my mentor. I thank her for her strong support on me at every phase of my doctoral study. Her meticulous scrutiny, prompt inspiration, scholarly advice, and scientific approach have enabled me to complete my doctoral degree. No words can fully express my gratitude to her. All the warm and touching moments were kept at the bottom of my heart.

Second, I am deeply thankful to my committee members, Dr. Shawna J. Lee, Dr. Vanessa Dalton, Dr. Marie-Anne Sanon Rosemberg, and Dr. James J. Yang. Their kind support and constructive and thoughtful feedback have helped me to a very great extent to complete my dissertation.

Third, I am extremely grateful to the Consulting for Statistics, Computing and Analytics Research (CSCAR) consultants for their support on the statistical analyses of my dissertation, especially the CSCAR director – Dr. Kerby Shedden. I appreciate all his great thoughts.

Fourth, I would like to express my gratitude to researchers who collected the National Health and Nutrition Examination Survey (NHANES) data as well as thousands of women that have contributed to our increasing understanding of allostatic load and its relationships with adverse perinatal outcomes.

Fifth, I would like to acknowledge the financial support from the Alice Hatt Lapidès Graduate Fellowship, the China Scholarship Council (CSC) Scholarship, and the Rackham One-Term Dissertation Fellowship.

Finally, I owe my deep sense of gratitude to my parents for their unceasing encouragement, support, and understanding throughout my doctoral study period. I also thank my friends for their constant encouragement and kind suggestions on my study and life.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	ii
LIST OF TABLES	vi
LIST OF FIGURES	vii
ABSTRACT	viii
Chapter 1: Introduction.....	1
Background and significance.....	1
Study aims	5
Justification on using the NHANES data.....	6
References	10
Chapter 2: Allostatic load: A theoretical model for understanding the relationship between maternal posttraumatic stress disorder and adverse birth outcomes.....	14
Abstract	14
Introduction	15
Methods	17
Theoretical overview.....	18
Empirical basis for main propositions.....	24
Explication of a new perinatal AL theoretical model	28
Evaluation of the new perinatal AL theoretical model	30
Implications for research.....	31
Potential implications for clinical practice, service delivery, and policy.....	32
Conclusion	33
References	35
Chapter 3: Exploring the optimal allostatic load scoring method in women of reproductive age.....	49
Abstract	49

Introduction.....	50
Methods.....	55
Results.....	60
Discussion.....	62
References.....	67
Chapter 4: Exploring the validity of allostatic load in pregnant women.....	88
Abstract.....	88
Introduction.....	89
Methods.....	91
Results.....	95
Discussion.....	96
References.....	103
Chapter 5: Conclusion	115
The summary of findings	115
Implications for research.....	118
Implications for clinical practice	119
Limitations and strengths	121
Future research directions	122

LIST OF TABLES

Table 1-1: Study aims	12
Table 2-1: Brief descriptions of the 4 theoretical models	42
Table 2-2: The new perinatal allostatic load model.....	44
Table 3-1: The descriptive statistics of sample sociodemographics and health outcomes (<i>N</i> =4319)	69
Table 3-2: The descriptive statistics of the 10 allostatic load indicators (<i>N</i> =4319)	70
Table 3-3: The descriptive statistics of allostatic load indices using the 5 scoring methods.....	71
Table 3-4: The descriptive statistics of tailored allostatic load indices (8 indicators) using the 5 scoring methods	72
Table 3-5: The binary logistic regressions of allostatic load indices by the 5 scoring methods on general health, diabetes, and hypertension.....	73
Table 3-6: The binary logistic regressions of tailored allostatic load indices by the 5 scoring methods on diabetes and hypertension.....	74
Table 3-7: The area under the ROC curve of allostatic load indices by the 5 scoring methods ...	75
Table 3-8: The area under the ROC curve of tailored allostatic load indices by the 5 scoring methods	76
Table 3-9: Evaluations of the 5 scoring methods	77
Table 3-10: The cut-off points, sensitivities, and specificities of the allostatic load index by the count-based method for predicting general health, diabetes, and hypertension.....	78
Table 4-1: Description of measures	106
Table 4-2: The descriptive statistics of demographics, gestational month, and history of adverse birth outcomes	107
Table 4-3: The summary of the 10 indicators in terms of their gestational patterns	108
Table 4-4: The associations between history of adverse birth outcomes and the allostatic load index.....	110
Table 4-5: The cut-off points, sensitivities, and specificities of the allostatic load index with history of adverse birth outcomes as the outcome.....	111

LIST OF FIGURES

Figure 2-1: The process from stress to poor health outcomes based on the allostatic load theory	45
Figure 2-2: The mediating role of allostatic load in the link between maternal PTSD and adverse birth outcomes	46
Figure 2-3: The influence of chronic stress, maternal risk behaviors, depression, and pregnancy-specific distress on the link between maternal PTSD and adverse birth outcomes.....	47
Figure 2-4: The theoretical model for understanding the link between maternal PTSD and adverse birth outcomes.....	48
Figure 3-1A: The 5 pure-type profiles with the probability of each allostatic load indicator at low or high risk for poor health.....	79
Figure 3-1B: The probability of each allostatic load indicator at low risk for poor health in the 5 pure-type profiles.....	80
Figure 3-2: The distributions of allostatic load indices using the 5 scoring methods	82
Figure 3-3: The distributions of tailored allostatic load indices using the 5 scoring methods.....	83
Figure 3-4: The ROC area statistics and 95% confidence intervals of each allostatic load summary measure in terms of general health, diabetes, and hypertension.....	84
Figure 3-5: The ROC area statistics and 95% confidence intervals of each tailored allostatic load summary measure in terms of diabetes and hypertension.....	85
Figure 3-6: The ROC curves of allostatic load indices by the 5 scoring methods for predicting general health, diabetes, and hypertension.....	86
Figure 3-7: The ROC curves of tailored allostatic load indices by the 5 scoring methods for predicting diabetes and hypertension.....	87
Figure 4-1: The adjusted means and standard errors of the allostatic load index by the count-based method across gestational month with allostatic load index scores in pregnant women with prior adverse birth outcomes plotted and the average allostatic load index in the non-pregnant population from the NHANES 2001-2006 data shown for comparison.....	112
Figure 4-2: The adjusted means and standard errors of the 10 allostatic load indicators across gestational month.....	114

ABSTRACT

Adverse perinatal outcomes such as preterm birth and low birth weight are significant public health concerns and contribute to neonatal morbidity and mortality. Maternal chronic stress (e.g., child maltreatment, posttraumatic stress disorder, depression) is an established predictor of adverse perinatal outcomes. However, the biological mechanisms by which maternal chronic stress affects adverse perinatal outcomes are less understood. Allostatic load (AL) refers to the cumulative dysregulations of multiple physiological systems responsive to multiple social-ecological levels of chronic stress. It is a promising conceptualization of the mechanism for stress effects on health. Little research has applied the AL theory to perinatal outcomes research to understand the complex pathophysiologic mechanisms for the stress-related adverse perinatal outcomes. Additionally, the optimal AL scoring method and the validity of pregnancy AL are less clear. Thus, the dissertation project had 3 aims: 1) to propose a theoretical model to situate AL in a role between maternal chronic stress and adverse perinatal outcomes; 2) to explore the optimal AL scoring method; and 3) to assess the gestational pattern of the AL summary score by the optimal scoring method and to test the validity of the pregnancy AL summary score for predicting a prior adverse birth outcome (as a proxy for adverse birth outcome subsequently on the current pregnancy). We used theory synthesis to construct a theoretical model to understand how maternal chronic stress contributes to adverse perinatal outcomes based on the AL theory. To address the second aim, women of reproductive age from the National Health and Nutrition Examination Survey (NHANES) data were included for analysis. We constructed AL summary

scores using 5 scoring methods including the count-based, Z-Score, logistic regression, factor analysis, and grade of membership method and validated each score. We found the ALI score by the logistic regression method had the best predictive performances with regard to general health status, diabetes, and hypertension, but differences among the 5 summary scores were minor. When the outcome information is known or consistent across different contexts, the logistic regression method is optimal for use; otherwise we recommended the count-based method. To address the third aim, pregnant women from the NHANES data were included for analysis. The ALI score at each gestational month was not different from the average ALI score ($M=2.35$, $SE=0.03$, $N=4319$) in the non-pregnant population, suggesting that measuring AL at any gestational time point would reflect women's true physiological functions as long as gestational age is considered when scoring AL. We also found poor predictive performance of the ALI score for predicting prior adverse birth outcomes, which suggested that the AL summary measure is not sufficiently sensitive to use as a single predictor for the risk of adverse birth outcomes. This dissertation project may lay theoretical and methodological underpinnings for future research to understand the etiologic contribution of maternal chronic stress to adverse perinatal outcomes. Empirical research on maternal chronic stress, AL, and perinatal outcomes would assist in identifying women at risk for adverse perinatal outcomes and developing and evaluating effective interventions to mitigate stress-related adverse perinatal outcomes.

Chapter 1: Introduction

Background and significance

Significance of studying adverse perinatal outcomes

Among adverse perinatal outcomes, preterm birth and low birth weight are significant contributors to U.S. health and socioeconomic disparities. In the United States, the preterm birth rate was around 10% in 2016 (Martin, Hamilton, & Osterman, 2017) and the low birth weight rate was 8% in 2013 (Matthews, MacDorman, & Thoma, 2015). Preterm birth and low birth weight have a substantial impact on infant death. In 2013, preterm birth accounted for 67% of infant deaths, which was the largest contributor to infant mortality. The infant mortality rate for infants born at less than 37 weeks of gestation was nearly 19 times the rate for infants born at full term (39-40 weeks of gestation). The infant mortality rate was highest for very preterm (less than 32 weeks) infants, 88 times the rate for full-term infants. The infant mortality rate for low birth weight (less than 2,500 grams) infants was 25 times higher than for infants born weighing 2,500 grams or more. Infants with birth weights of less than 1,000 grams accounted for nearly one-half of infant mortality in the United States (Matthews et al., 2015). Premature or low birth weight infants also had a much greater risk of neurodevelopmental impairments as children (e.g., low intelligence quotient and cerebral palsy; Allen, 2008; Goldenberg, Culhane, Iams, & Romero, 2008; Green et al., 2005) as well as all-cause mortality as adults (Risnes et al., 2011). Furthermore, those adverse perinatal outcomes cause a long-lasting financial burden on the entire

society. In 2005, the Institute of Medicine reported that the annual cost associated with preterm birth in the United States was more than 26 billion dollars (Butler, & Behrman, 2007).

However, due to the multifactorial nature of preterm birth, the mechanisms that contribute to preterm birth have been poorly understood. More research needs to be directed toward identifying the potential pathways for the risk of preterm birth. Until multiple pathways leading to preterm birth are more clearly understood, effective risk identification and interventions could be developed to reduce the occurrence of preterm birth.

The mechanisms leading to adverse perinatal outcomes

The causes of preterm birth are numerous and complex, including mechanical and medical factors (i.e., multiple gestation, uterine and cervical abnormalities, placental abruption, vaginal bleeding, fetal defects) and infection/inflammation (i.e., urogenital, pneumonia, sexually transmitted, periodontal disease; Latendresse, 2009). Increasing attention has been also paid to maternal chronic stress including socioeconomic status, child maltreatment, intimate partner violence, depression, posttraumatic stress disorder (PTSD), perceived stress, etc.. The associations between maternal chronic stress and preterm birth have been well documented in prior studies (Hellgren, Akerud, Skalkidou, & Sundstrom-Poromaa, 2013; King et al., 2010; Shaw et al., 2014; Shea et al., 2007; Voegtline et al., 2013). However, the mechanisms accounting for the impact of maternal chronic stress on preterm birth are much less clear.

According to the 2005 research agenda for preterm birth, maternal chronic stress and its biological pathways may be a major contributor to preterm birth (Green et al., 2005). Maternal chronic stress may activate the maternal/fetal hypothalamic-pituitary-adrenal (HPA) axis, enhance maternal/fetal/intrauterine inflammatory processes, and reduce both uterine and umbilical blood flow. This cascade of physiologic events occurs in the mother, uterus, placenta

and fetus and interacts with maternal and fetal physiology during pregnancy, which eventually initiates the premature labor process (Latendresse, 2009).

Maternal and fetal stress responses and the cascade of physiologic events could serve as a potential pathophysiologic mechanism for stress-related preterm birth. Given the complexity of the physiologic stress reactivity and preterm birth, the theory of allostatic load (AL) may assist in better understanding the associations between maternal chronic stress and adverse birth outcomes. Because of biomedical limitations, psychosocial care to address maternal stress has been recommended as a strategy to improve mother and infant health (Renfrew et al., 2014). Thus, improved understanding of the impact of maternal chronic stress on adverse birth outcomes may contribute to effective psychosocial care.

AL as a potential mechanism for the impact of maternal chronic stress on adverse perinatal outcomes

AL is the accumulated physiological dysregulations across multiple body systems including neuroendocrine, immune, cardiovascular, and metabolic systems resulting from repeated, chronic stress. In response to chronic stress, multiple physiological systems are activated and activation is measurable in biomarkers (i.e., cortisol, C-reactive protein, cholesterol, glucose). The overexpression of these biomarkers can have a substantial impact on the body systems, leading to poor health outcomes (McEwen, 2006). The AL theory explains the mechanism of stress effects on health. Applying the AL theory to perinatal outcomes research may enhance our understanding of the relationships between maternal chronic stress and adverse perinatal outcomes. AL is commonly operationalized as a single summary measure by combining multiple physiological indicators from different systems. It may be theoretically posited that AL mediates the relationships between maternal chronic stress and adverse perinatal outcomes. A

theoretical model based on the AL theory is needed to delineate the complex biological mechanisms by which maternal chronic stress leads to adverse perinatal outcomes, which may be used to guide future perinatal research.

It is crucial to elucidate the underlying mechanisms by which maternal chronic stress affects preterm birth. The enhanced understanding of these pathways could allow perinatal researchers and health care providers to develop risk-screening approaches for early detection of high-risk women as well as design effective preventions and interventions to reduce the high rates of premature and low birth weight infants and related infant mortality and morbidity.

Methodological issues of measuring AL

A few methodological issues need to be addressed before investigating AL as a pathway for understanding the effects of maternal chronic stress on adverse perinatal outcomes. The first issue is how to best score AL. AL is operationalized as a single index that combines multiple physiological indicators (the result of a lab assay of a biomarker from immune, cardiovascular, or metabolic systems or anthropometric measure) from different systems. Varied scoring approaches, including the count-based, Z-Score, canonical correlation, recursive partitioning, grade of membership, etc., have been used to score AL in previous studies. Those scoring approaches have their own strengths and weaknesses. It is unknown which method is optimal to summarize a complex set of biological measurements in an interpretable way. No studies have focused on comparing different scoring methods and making recommendations for the optimal scoring approach. The second issue is that the validity of measuring AL in pregnancy is unclear. There are great physiological changes during pregnancy, with altered physiological stress reactivity. It is unclear whether the alterations in AL-related indicators during pregnancy could reflect women's true AL levels. The last issue is that the predictive validity of pregnancy AL for

predicting adverse perinatal outcomes is unclear. While prior studies examined the predictive validity of AL for predicting self-rated health and physical and cognitive performance (Karlamañla, Singer, McEwen, Rowe, & Seeman, 2002; Seplaki, Goldman, Gleib, & Weinstein, 2005), investigations of the predictive performance of pregnancy AL for predicting adverse perinatal outcomes are lacking. Various risk-scoring systems and the 2 technical assessments (i.e., ultrasound examination of cervical length and fetal fibronectin screening) have reported a low accuracy for predicting preterm birth (DeFranco, Lewis, & Odibo, 2013; Honest et al., 2004). Those screening tools could not identify most women who subsequently have an adverse perinatal outcome. Thus, more work is definitely needed to explore other effective risk indicators that could identify the majority of women subsequently having adverse perinatal outcomes. The AL summary score may be a potential “early warning” sign for adverse perinatal outcomes.

Study aims

This dissertation project includes three parts that address each of the research questions listed in Table 1-1.

Describing the effects of maternal chronic stress on adverse perinatal outcomes and their biological pathways within a theoretical model may provide a better understanding of the etiology of stress-related adverse perinatal outcomes. Although the whole theoretical model cannot be tested at this point due to data limitations, the goal of the dissertation project is to elaborate a theoretical model for future perinatal outcomes research and to focus on a few methodological issues of measuring AL as an initial step toward future research. This dissertation project may provide a theoretical and methodological groundwork for understanding the etiologic contribution of maternal chronic stress to adverse perinatal outcomes. Empirical research on the associations between maternal chronic stress, AL, and adverse perinatal

outcomes would assist in identifying women at risk for adverse perinatal outcomes and developing and evaluating clinical interventions early in gestation that might modify maternal perceptions or experiences of stress and their impact on adverse perinatal outcomes. Since certain maternal chronic stress such as depression and PTSD is potentially modifiable, preventable, or treatable, there are also implications for public policy and population health. With further research on demonstrating AL as a mediator in the associations between maternal chronic stress and adverse perinatal outcomes, AL may serve as an objective screening assessment for both maternal chronic stress and adverse perinatal outcomes.

Justification on using the NHANES data

The second and third parts of the dissertation that aimed to address a few methodological issues were secondary analyses of data from the National Health and Nutrition Examination Survey (NHANES). NHANES is a cross-sectional study with a complex, multistage probability sampling design used to select a sample representative of the civilian non-institutionalized resident population of the United States, which has been conducted in 2-year cycles since 1999. The annual sample size has been approximately 5,000 individuals from 15 different locations (12 locations for 1999) selected from all 50 states and the District of Columbia. The survey consists of questionnaires administered in the home, followed by a standardized physical examination in a specially designed and equipped mobile examination center (MEC). The examination component of NHANES includes medical, dental, and physiological measurements, as well as numerous laboratory tests to assess various aspects of health.

The NHANES database has several advantages for the second and third aims of the dissertation. First, the NHANES database includes numerous physiological measures from multiple body systems, which provides adequate AL indicators for our study to create AL index

(ALI). Second, the sample of the NHANES includes pregnant women. The sample size of pregnant women is large enough to construct gestational-age-specific ALI with many multi-system physiological measures. None of other available public databases has pregnancy physiological measures and perinatal outcomes of interest that can be used to create pregnancy-specific ALI and validate the index. Finally, data are collected with standardized procedures and protocols to assure that the data for these analyses are of high quality in terms of reliability and validity.

The third part of the dissertation was a preliminary study that aimed to assess the predictive performance of pregnancy ALI scores for predicting adverse perinatal outcomes. Because of the cross-sectional study design of the NHANES, the analysis of physiological AL indicators from pregnant women in the sample cannot be done using the outcomes of the current pregnancy as the dependent variable. Information on subsequent perinatal outcomes for pregnant participants is unknown for addressing the third aim of the dissertation. Only history of delivering premature low birth weight infants has been measured. Thus, history of delivering premature low birth weight infants was used as the outcome variable to address the third aim of the dissertation. Although prospective data would be better, cross-sectional studies are often used in early phases of research on a topic. These cross-sectional data have many strengths for the methodological focus of this work. And there are a few reasons why using the woman's previous birth outcomes as a proxy of her risk for adverse birth outcomes in this pregnancy is valid and reasonable.

First, the literature supports the view that prior adverse birth outcome is a risk factor for a subsequent outcome. Women who reported history of delivering low birth weight infants are at high risk for subsequently delivering low birth weight infants. Using the 2002 birth certificate

data from New Jersey and Missouri, among women with a prior spontaneous preterm birth (<37 weeks), the probability of having a preterm birth in the subsequent pregnancy was 22.5%. For those who reported a history of very preterm birth (<32 weeks), the chance of preterm birth recurrence is 33% (Petrini et al., 2005). The risk of recurrent preterm birth increases as the number of previous preterm births increases, and thus is highest in women with more than 1 previous preterm delivery (Esplin et al., 2008; McManemy, Cooke, Amon, & Leet, 2007). In a retrospective cohort study of 3334 women having two births in Washington State between 1984 and 1991, women who delivered a very low birth weight first infant were at an 11.5-fold increased risk of delivering a low birth weight (less than 2500 g) second infant (Bratton, Shoultz, & Williams, 1996). A history of low birth weight in previous pregnancies is the strongest predictor of low birth weight delivery in a subsequent pregnancy, and this risk is maintained after controlling for different demographic and obstetric characteristics (Raine, Powell, & Krohn, 1994). Similarly, in a cohort study of 10,397 pregnant women enrolled at seven medical centers between 1984 and 1989, delivering a low-birth-weight infant in a woman's prior pregnancy was most strongly associated with delivering a premature infant in the current pregnancy (Hillier et al., 1995).

Second, this is recognized clinically. One of recommendations made by the Advisory Committee on Infant Mortality (2001) for reducing the low birth weight rate was to develop societal and governmental policies for providing family planning services especially for women with prior low birth weight deliveries.

Third, a published case-control study using the NHANES data already used "history of small for gestational age or preterm birth" as an independent variable. That study's purpose was to assess the association of non-pregnancy AL with birth outcomes. They demonstrated that non-

pregnant women with history of small for gestational age or preterm birth infants had higher AL levels compared to those with normal birth outcomes (Hux, Catov, & Roberts, 2014). They did not limit the sample to currently pregnant women and they did not have our methodological focus. But the significant association between non-pregnancy AL and prior adverse birth outcomes provides evidence for the third part of the dissertation to test the predictive validity of pregnancy AL for predicting a proxy for adverse birth outcomes (history of adverse birth outcomes).

Thus, the third part of the dissertation which is methodological in focus, tested the validity of an ALI created from indicators collected in pregnancy and built upon positive findings from an already-published project. Higher AL levels derived during pregnancy may also prove to be detectable in this high-risk population. Therefore, to address the third aim of the dissertation, we used the “prior adverse birth outcome” as the outcome variable and framed it as “risk for future adverse birth outcome”. Although the cross-sectional study design of the NHANES is a limitation, it is acceptable for the methodological purpose of the study.

Taken together, this dissertation constructed a theoretical model to situate AL in a mediating role between maternal chronic stress and adverse perinatal outcomes and did secondary analysis using a large national data to address issues regarding how to best score AL and the validity of AL in pregnancy. It may lay a theoretical and methodological groundwork for future perinatal outcomes research.

References

- Advisory Committee on Infant Mortality (2001). Low birth weight report and recommendations. <http://www.hrsa.gov/advisorycommittees/mchbadvisory/InfantMortality/About/lowbirthweight.pdf>.
- Butler, A. S., & Behrman, R. E. (Eds.). (2007). *Preterm birth: causes, consequences, and prevention*. National Academies Press.
- Allen, M. C. (2008). Neurodevelopmental outcomes of preterm infants. *Current Opinion in Neurology*, 21(2), 123-128. doi: 10.1097/WCO.0b013e3282f88bb4
- Bratton, S. L., Shoultz, D. A., & Williams, M. A. (1996). Recurrence risk of low birthweight deliveries among women with a prior very low birthweight delivery. *American Journal of Perinatology*, 13(3), 147-150. doi: Doi 10.1055/S-2007-994313
- DeFranco, E. A., Lewis, D. F., & Odibo, A. O. (2013). Improving the screening accuracy for preterm labor: is the combination of fetal fibronectin and cervical length in symptomatic patients a useful predictor of preterm birth? A systematic review. *American Journal of Obstetrics & Gynecology*, 208(3). doi: ARTN 233.e110.1016/j.ajog.2012.12.015
- Esplin, M. S., O'Brien, E., Fraser, A., Kerber, R. A., Clark, E., Simonsen, S. E., . . . Varner, M. W. (2008). Estimating recurrence of spontaneous preterm delivery. *Obstetrics & Gynecology*, 112(3), 516-523. doi: Doi 10.1097/Aog.0b013e318184181a
- Goldenberg, R. L., Culhane, J. F., Iams, J. D., & Romero, R. (2008). Epidemiology and causes of preterm birth. *Lancet*, 371(9606), 75-84. doi: 10.1016/S0140-6736(08)60074-4
- Green, N. S., Damus, K., Simpson, J. L., Iams, J., Reece, E. A., Hobel, C. J., . . . March Of Dimes Scientific Advisory Committee On, P. (2005). Research agenda for preterm birth: recommendations from the March of Dimes. *American Journal of Obstetrics & Gynecology*, 193(3 Pt 1), 626-635. doi: 10.1016/j.ajog.2005.02.106
- Hellgren, C., Akerud, H., Skalkidou, A., & Sundstrom-Poromaa, I. (2013). Cortisol awakening response in late pregnancy in women with previous or ongoing depression. *Psychoneuroendocrinology*, 38(12), 3150-3154. doi: 10.1016/j.psyneuen.2013.08.007
- Hillier, S. L., Nugent, R. P., Eschenbach, D. A., Krohn, M. A., Gibbs, R. S., Martin, D. H., . . . Klebanoff, M. A. (1995). Association between bacterial vaginosis and preterm delivery of a low-birth-weight infant. *New England Journal of Medicine*, 333(26), 1737-1742. doi: Doi 10.1056/Nejm199512283332604
- Honest, H., Bachmann, L. M., Sundaram, R., Gupta, J. K., Kleijnen, J., & Khan, K. S. (2004). The accuracy of risk scores in predicting preterm birth--a systematic review. *Journal of Obstetrics and Gynaecology*, 24(4), 343-359. doi: 10.1080/01443610410001685439
- Hux, V. J., Catov, J. M., & Roberts, J. M. (2014). Allostatic load in women with a history of low birth weight infants: the national health and nutrition examination survey. *Journal of Women's Health (Larchmt)*, 23(12), 1039-1045. doi: 10.1089/jwh.2013.4572
- Karlamangla, A. S., Singer, B. H., McEwen, B. S., Rowe, J. W., & Seeman, T. E. (2002). Allostatic load as a predictor of functional decline MacArthur studies of successful aging. *Journal of Clinical Epidemiology*, 55(7), 696-710. doi: Pii S0895-4356(02)00399-2

- King, N. M., Chambers, J., O'Donnell, K., Jayaweera, S. R., Williamson, C., & Glover, V. A. (2010). Anxiety, depression and saliva cortisol in women with a medical disorder during pregnancy. *Archives of Women's Mental Health*, 13(4), 339-345. doi: 10.1007/s00737-009-0139-5
- Latendresse, G. (2009). The Interaction Between Chronic Stress and Pregnancy: Preterm Birth from A Biobehavioral Perspective. *Journal of Midwifery & Womens Health*, 54(1), 8-17. doi: 10.1016/j.jmwh.2008.08.001
- Martin, J. A., Hamilton, B. E., & Osterman, M. J. K. (2017). Births in the United States, 2016. *NCHS Data Brief*, (287), 1-8.
- Matthews, T. J., MacDorman, M. F., & Thoma, M. E. (2015). Infant Mortality Statistics From the 2013 Period Linked Birth/Infant Death Data Set. *National Vital Statistics Reports*, 64(9), 1-30.
- McEwen, B. S. (2006). Protective and damaging effects of stress mediators: central role of the brain. *Dialogues in Clinical Neuroscience*, 8(4), 367-381.
- McManemy, J., Cooke, E., Amon, E., & Leet, T. (2007). Recurrence risk for preterm delivery. *American Journal of Obstetrics and Gynecology*, 196(6), 576-577. doi: ARTN 576.e110.1016/j.ajog.2007.01.039
- Petrini, J. R., Callaghan, W. M., Klebanoff, M., Green, N. S., Lackritz, E. M., Howse, J. L., . . . Damus, K. (2005). Estimated effect of 17 alpha-hydroxyprogesterone caproate on preterm birth in the United States. *Obstetrics & Gynecology*, 105(2), 267-272. doi: 10.1097/01.AOG.0000150560.24297.4f
- Raine, T., Powell, S., & Krohn, M. A. (1994). The Risk of Repeating Low-Birth-Weight and the Role of Prenatal-Care. *Obstetrics & Gynecology*, 84(4), 485-489.
- Renfrew, M. J., McFadden, A., Bastos, M. H., Campbell, J., Channon, A. A., Cheung, N. F., . . . Declercq, E. (2014). Midwifery and quality care: findings from a new evidence-informed framework for maternal and newborn care. *Lancet*, 384(9948), 1129-1145. doi: 10.1016/S0140-6736(14)60789-3
- Risnes, K. R., Vatten, L. J., Baker, J. L., Jameson, K., Sovio, U., Kajantie, E., . . . Bracken, M. B. (2011). Birthweight and mortality in adulthood: a systematic review and meta-analysis. *International Journal of Epidemiology*, 40(3), 647-661. doi: 10.1093/ije/dyq267
- Seplaki, C. L., Goldman, N., Gleib, D., & Weinstein, M. (2005). A comparative analysis of measurement approaches for physiological dysregulation in an older population. *Experimental Gerontology*, 40(5), 438-449. doi: 10.1016/j.exger.2005.03.002
- Shaw, J. G., Asch, S. M., Kimerling, R., Frayne, S. M., Shaw, K. A., & Phibbs, C. S. (2014). Posttraumatic stress disorder and risk of spontaneous preterm birth. *Obstetrics & Gynecology*, 124(6), 1111-1119. doi: 10.1097/AOG.0000000000000542
- Shea, A. K., Streiner, D. L., Fleming, A., Kamath, M. V., Broad, K., & Steiner, M. (2007). The effect of depression, anxiety and early life trauma on the cortisol awakening response during pregnancy: Preliminary results. *Psychoneuroendocrinology*, 32(8-10), 1013-1020. doi: 10.1016/j.psyneuen.2007.07.006
- Voegtline, K. M., Costigan, K. A., Kivlighan, K. T., Laudenslager, M. L., Henderson, J. L., & DiPietro, J. A. (2013). Concurrent levels of maternal salivary cortisol are unrelated to self-reported psychological measures in low-risk pregnant women. *Archives of Women's Mental Health*, 16(2), 101-108. doi: 10.1007/s00737-012-0321-z
- Walker, L. O., & Avant, K. C. (2005). Strategies for theory construction in nursing.

Table 1-1: Study aims

	Aims	Methods
Study 1: A theoretical model construction	To propose a theoretical model to guide research on the mediating role of AL in the associations between maternal chronic stress and adverse perinatal outcomes.	We used theory synthesis as a strategy of theory building (Walker & Avant, 2005) and followed these procedures: (1) specifying focal concepts; (2) identifying related factors and relationships; and (3) constructing an integrated representation.
Study 2: A secondary data analysis using the NHANES database	To determine the optimal AL scoring method in women of reproductive age.	<ul style="list-style-type: none"> • Women of reproductive age from the NHANES 2001-2006 cycles were included. • The count-based, Z-Score, multivariable logistic regression, factor analysis, and the grade of membership method were used to create ALI. • Binominal logistic regressions were conducted with the ALI score as the independent variable and general health, diabetes, and hypertension as the dependent variables to test the strength of the associations between ALI scores and health outcomes. • The area under the ROC curve statistics was also calculated to examine the predictive validity of the ALI by each scoring approach for predicting each of the three health outcomes. • The strengths and weaknesses of each scoring approach were qualitatively compared to make recommendation.
Study 3: A secondary data analysis using the NHANES database	<ul style="list-style-type: none"> • To examine the gestational pattern of AL indicators and the ALI and to test whether the ALI has face validity in pregnancy despite physiological changes of pregnancy; • To assess the predictive validity of pregnancy ALI on a proxy for adverse birth outcomes, having had a premature low birth 	<ul style="list-style-type: none"> • Pregnant women from the NHANES 1999-2006 cycles were included. • The ALI score was created by the optimal scoring method recommended from Study 2. • Curves were plotted to describe changes in the ALI score and each individual indicator across different gestational month.

	weight infant on the previous pregnancy.	<ul style="list-style-type: none"> • Binomial logistic regression models with the ALI score as the independent variable and history of adverse birth outcomes as the dependent variable were conducted to examine the associations between the ALI score and history of adverse birth outcomes. • The cut-off points, sensitivities, and specificities of the ALI score were also computed.
--	--	---

Note. NHANES, National Health and Nutritional Examination Survey; AL, allostatic load; ALI, allostatic load index.

Chapter 2: Allostatic load: A theoretical model for understanding the relationship between maternal posttraumatic stress disorder and adverse birth outcomes

Abstract

Adverse birth outcomes such as preterm birth and low birth weight are significant public health concerns and contribute to neonatal morbidity and mortality. Studies have increasingly been exploring the predictive effects of maternal posttraumatic stress disorder (PTSD) on adverse birth outcomes. However, the biological mechanisms by which maternal PTSD affects birth outcomes are not well understood. Allostatic load refers to the cumulative dysregulations of the multiple physiological systems as a response to multiple social-ecological levels of chronic stress. Allostatic load has been well documented in relation to both chronic stress and adverse health outcomes in the non-pregnant population. However, the mediating role of allostatic load is less understood when it comes to maternal PTSD and adverse birth outcomes. The purpose of the study is to propose a theoretical model that depicts how allostatic load could mediate the impact of maternal PTSD on birth outcomes. We followed the procedures for theory synthesis approach described by Walker and Avant (2011), including specifying focal concepts, identifying related factors and relationships, and constructing an integrated representation. We first present a theoretical overview of the allostatic load theory and the other 4 relevant theoretical models. Then we provided a brief narrative review of literature that empirically supports the propositions of the integrated model. Finally, we described our theoretical model. The theoretical model synthesized has the potential to advance perinatal research by delineating multiple biomarkers to

be used in future. After it is well validated, it could be utilized as the theoretical basis for health care professionals to identify high-risk women by evaluating their experiences of psychosocial and traumatic stress and to develop and evaluate service delivery and clinical interventions that might modify maternal perceptions or experiences of stress and eliminate their impacts on adverse birth outcomes.

Keywords: posttraumatic stress disorder, preterm birth, low birth weight, allostatic load, theoretical model

Introduction

Adverse birth outcomes including preterm birth and low birth weight are significant public health concerns and contribute to neonatal morbidity and mortality (Lawn et al., 2010; Lawn, Wilczynska-Ketende, & Cousens, 2006). In 2013, the rates for preterm birth and low birth weight were 11.4% and 8% respectively in the United States (Martin, Hamilton, Osterman, Curtin, & Matthews, 2015). Two-thirds of low birth weight (infants born <2500g) co-occurs with preterm birth (Martin et al., 2012). Preterm birth has been estimated to cause up to 50% of children's neurodevelopment problems (e.g., low intelligence quotient, and cerebral palsy; Allen, 2008; Goldenberg, Culhane, Iams, & Romero, 2008; Green et al., 2005; Latendresse, 2009). Furthermore, a systematic review and meta-analysis suggested an inverse association of birth weight with all-cause mortality in adults (Risnes et al., 2011).

A growing body of perinatal research has focused on maternal psychosocial and sociodemographic stress and depression and their biological pathways leading to adverse birth outcomes (Hellgren, Akerud, Skalkidou, & Sundstrom-Poromaa, 2013; King et al., 2010; Shea et al., 2007; Voegtline et al., 2013). However, traumatic stress and its posttraumatic stress disorder (PTSD) sequelae have been much less frequently studied in relation to adverse birth outcomes.

In a study conducted in a large national sample of 9282 individuals aged 18 years and older from the United States, the overall estimated lifetime prevalence of PTSD was 6.8% (Kessler et al., 2005a) and lifetime prevalence among females was 12.3% in the U.S. National Women's Study (Resnick, Kilpatrick, Dansky, Saunders, & Best, 1993). A recent systematic review showed the mean prevalence of PTSD in pregnancy was 3.3% (95%CI: 2.44-4.54) in community samples and 18.95% (95%CI: 10.62-31.43) in high risk samples (e.g., women with nausea, vomiting, or hyperemesis, or a history of child maltreatment) and the average rate for comorbidity of PTSD and depression was 44.1% in pregnancy in 6 studies (Yildiz, Ayers, & Phillips, 2017). Thus, considering the impact of maternal pre-existing and current trauma exposure and PTSD may add better understanding of the stress-related etiology of adverse birth outcomes than focus on depression alone.

Prior studies reported inconsistent findings with regard to the associations between maternal PTSD and birth outcomes, which may be due to small sample sizes, lack of generalizability, and failing to controlling for some critical confounding factors including the type of trauma (e.g., child maltreatment, and intimate partner violence), demographic stress (e.g., socioeconomic status), co-occurring risk behaviors (e.g., tobacco, and drug use), and co-morbidities (e.g., depression) of PTSD (Berkowitz et al., 2003; Chang, Chang, Lin, & Kuo, 2002; Engel, Berkowitz, Wolff, & Yehuda, 2005; Lipkind, Curry Ae Fau - Huynh, Huynh M Fau - Thorpe, Thorpe Le Fau - Matte, & Matte, 2010; Morland et al., 2007; Rogal et al., 2007; Rosen, Seng Js Fau - Tolman, Tolman Rm Fau - Mallinger, & Mallinger, 2007; Seng et al., 2001; Xiong et al., 2008). The study by Seng, Low, Sperlich, Ronis, & Liberzon (2011b) and the two relatively recent studies (Shaw et al., 2014; Yonkers et al., 2014) addressed several limitations of

these previous studies and consistently found significant relationships between maternal PTSD and adverse birth outcomes.

Although significant relationships between maternal PTSD and adverse birth outcomes were detected by these more recent studies with large representative samples (Seng et al., 2011b; Shaw et al., 2014; Yonkers et al., 2014), it is unclear what plausible biological pathways are involved via which maternal PTSD affects birth outcomes. Describing the impact of maternal lifetime PTSD on adverse birth outcomes and its biological pathways within the context of a theoretical model could suggest new ways for understanding the etiology of adverse birth outcomes.

Methods

We use theory synthesis as a strategy of theory building (Walker and Avant, 2011). It offers a way to synthesize existing theories and empirical evidence. The steps involved in theory synthesis include: (1) specifying focal concepts; (2) identifying related factors and relationships; and (3) constructing an integrated representation. First, to identify an applicable conceptualization of the relationship between stress and health and its neurobiological mechanisms, we undertook a broad literature search on PubMed and google scholar with “stress/traumatic stress/trauma”, “health/health outcome/disease”, “biology”, and “theory” as search terms. The theory of allostatic load (AL; McEwen and Stellar, 1993) was chosen because it is being applied in relation to stress broadly defined to include toxic and traumatic stress and thus is useful for PTSD research (Shonkoff et al., 2012). AL serves as a strong theoretical basis for understanding the complex and multiple biological mechanisms of the impact of maternal PTSD on adverse birth outcomes. Based on the AL theory, we specified focal concepts that could be included in the synthesized theoretical model. We also reviewed and drew upon 4

previous theoretical models to build our proposed model. Then, to identify variables related to the focal concepts as well as the relationships among variables, we did another broad literature search on PubMed and google scholar with “posttraumatic stress disorder”, “child maltreatment/childhood trauma”, “stress/traumatic stress/trauma”, “depression”, “pregnancy”, “allostatic load”, and “preterm birth/low birth weight/pregnancy complication” as search terms. Finally, we organized the concepts into an integrated network and employed a diagram to holistically depict interrelationships among concepts.

This article was organized as follows. We first present a theoretical overview of the AL theory. Then we synthesize the work of 4 authors (Beckie, 2012; Olson et al., 2015; Premji, 2014; Seng, 2002) whose theoretical models contribute to the model we propose. Synthesis of the contributions of the four models and the additional contributions of our model are then discussed. A brief narrative review that empirically supports the propositions of the integrated model is then provided. Finally, our theoretical model is described and implications for research, clinical practice, service delivery, and policy are discussed.

Theoretical overview

The AL theory

The AL theory is a lifespan theory that may provide a theoretical basis for understanding the relationships between maternal PTSD and birth outcomes. AL refers to the accumulated multi-system physiologic dysfunction resulting from chronic or severe stress that could ultimately lead to disease (McEwen, 1998). When stress (e.g., traumatic life events, child maltreatment) occurs, physiological mediators (e.g., glucocorticoid hormones) are produced from multiple systems (e.g., neuroendocrine, immune, and cardiovascular systems) to generate the physiological adaptations to stress (McEwen and Wingfield, 2003). The process from stress

to poor health outcomes includes three stages: primary mediators, secondary outcomes, and tertiary outcomes (McEwen and Seeman, 1999). First, the neuroendocrine system is activated and primary mediators (i.e., cortisol, epinephrine, norepinephrine) are produced to generate the acute stress response. Second, the other systems produce physiological changes to generate a more long-term stress response, which leads to secondary outcomes including dysregulations in immune (i.e., interleukin-6, C-reactive protein), metabolic (i.e., high density lipoprotein, glycosylated hemoglobin, glucose), and cardiovascular (i.e., systolic blood pressure, diastolic blood pressure) systems, and changes in anthropometric parameters (i.e., body mass index, waist/hip). Finally, the accumulated physiological dysregulations across multiple systems result in the tertiary outcomes with clinical manifestation including morbidity and mortality (Figure 2-1). AL includes primary mediators and secondary outcomes. In response to chronic or severe stress, these physiological mediators remain at abnormal levels. The overexpression of these mediators has damaging effects on the body systems, leading to diseases (e.g., cardiovascular diseases, diabetes; McEwen, 2006).

This cumulative, multi-system framework provides a more significant and comprehensive pathway for understanding physiological predictors of health risk than the more common approach of concentrating on risks associated with individual systems. A composite measure of biomarkers may be a stronger predictor of adverse birth outcomes in women with PTSD (Seeman, McEwen, Rowe, & Singer, 2001). AL is commonly operationalized by combining multiple physiological indicators from different systems into one single summary score. The count-based method has been the most frequently used method, but other scoring methods such as the Z-Score, canonical correlation, and grade of membership method have been also used to create an AL index. Few studies have compared those scoring methods and thus little is known

which method is optimal for use. Additionally, many combinations of biomarkers and anthropometric indicators have been included in the AL summary measure. More work is needed to determine the optimal scoring method and the best combinations of physiological indicators so that the AL summary measure could reliably reflect the multi-system physiological dysregulations.

Despite those methodological issues regarding measuring AL, a substantial number of studies have examined AL levels in relation to stressors or health outcomes using data from large national-level studies including the MacArthur Study of Successful Aging (MSSA), the Taiwanese Social Environment and Biomarkers of Aging Study (SEBAS), and the National Health and Nutrition Examination Survey (NHANES). Lower SES and higher social adversity were found to correlate with higher AL summary score (Chyu and Upchurch, 2011; Gustafsson, Janlert, Theorell, Westerlund, & Hammarstrom, 2011, 2012; Seeman et al., 2008). In the MacArthur cohort study with a sample of 70- to 79-year-old American adults, higher baseline AL summary score was highly correlated with cardiovascular diseases (CVD), cognitive and physical decline, and all-cause mortality over 12 years of follow-up (Gruenewald, Seeman, Karlamangla, & Sarkisian, 2009; Seeman et al., 2004; Seeman et al., 2001; Seeman, Singer, Rowe, Horwitz, & McEwen, 1997). Higher AL summary score was also associated with increased frailty 3 years later in these elders (Gruenewald et al., 2009). About 35.4% of SES effect on mortality was mediated by the AL index (Seeman et al., 2004). In the SEBAS, increased AL summary score was associated with poorer health status (e.g., self-rated health, activities of daily living, mobility) and cognitive impairments (Seplaki, Goldman, Weinstein, & Lin, 2006). Higher AL summary score was also associated with increased 3-year mortality risk (Goldman, Turra, Gleib, Lin, & Weinstein, 2006). Cross-sectional studies using the NHANES

data examined the association between AL and all-cause mortality and found higher AL summary score in relation to increased risk for mortality (Borrell, Dallo, & Nguyen, 2010; Crimmins, Kim, & Seeman, 2009). Participants with an AL summary score of 2 and ≥ 3 had mortality rates that were 40% and 88% greater than those with a score of ≤ 1 (Borrell et al., 2010). Taken together, AL has been reported in relation to psychosocial and sociodemographic stress as well as various health outcomes. However, AL remains to be fully applied in the context of women of childbearing age and related health outcomes (Rosemberg et al., 2017).

According to the 2005 March of Dimes Research Agenda for Preterm Birth, preterm birth is conceptualized as a “common, complex disorder” caused by multiple pathways that involve interactions among genetic, environmental, social, and behavioral factors. Maternal stress and stress-related physiological responses are one of 5 major pathways to preterm birth (Green et al., 2005). Maternal psychosocial, emotional, or other environmental stress could increase the risk of developing preterm birth through several interacting physiological pathways: (1) neuroendocrine mechanisms via activating the activities of maternal/fetal hypothalamic-pituitary-adrenal (HPA) axes that could facilitate premature labor and enhance placental-fetal endocrine activities; (2) immune pathways via maternal/fetal/intrauterine inflammatory processes, resulting in vulnerability to maternal infection and promoting preterm birth; and (3) maternal-placental-fetal vascular mechanisms via the reduction of both uterine and umbilical blood flow. Maternal stress causes the accumulated physiological changes from multiple systems including neuroendocrine, immune, and cardiovascular and eventually leads to preterm birth, which is consistent with the AL theory (McEwen, 2000). Additionally, given that immune/inflammatory, abnormal uterine distention, bleeding/thrombophilias, and toxins or hormones are the other 4 pathways to preterm birth, the cardiovascular and immune/inflammatory dysregulations that are parts of AL may also

contribute some extent of stress-related contributions to the immune/inflammatory and bleeding/thrombophilias pathways to prematurity.

Synthesis of 4 previous theoretical models contributing to the model we propose

Beckie (2012) proposed a heuristic AL model for understanding cumulative, multisystem physiological consequences of health disparities that could ultimately contribute to diseases. In that model, allostatic challenges, including child adversity and genetic, environmental, sociodemographic, psychosocial, behavioral, and clinical challenges, initiate the AL process involving changes in primary mediators, secondary outcomes, and tertiary outcomes. According to Beckie's model, primary mediators include physiological indicators from neuroendocrine and immune/inflammatory systems, secondary outcomes are systemic dysregulations of metabolic, cardiovascular, and inflammatory biomarkers, and tertiary outcomes of AL emerge with clinical manifestations such as physical and mental health problems, quality of life, and mortality. The model does not focus on any specific health outcome and is not pregnancy-specific, but provides us an organizing, cumulative, and multisystem framework that can be operationalized in research studies and can be a theoretical basis for proposing a more specific model within the childbearing context.

Seng's (2002) conceptual framework focused on the impact of maternal PTSD on perinatal outcomes. It posited that PTSD's effect might be mediated by behavioral and neuroendocrine alterations. But it did not consider the accumulated dysregulations from multiple physiological systems that result from maternal traumatic stress and PTSD and contribute to adverse perinatal outcomes. The framework was not informed by the AL theory.

Olson et al. (2015) and Premji (2014) used the AL theory and constructed theoretical models to link perinatal distress or cumulative life stressors to birth outcomes through AL.

Premji (2014) proposed a theoretical model based on the AL theory to understand the contribution of multiple forms of maternal perinatal distress on pregnancy outcomes for women in low- and middle-income countries. In that theoretical model, AL was considered as a mediating pathway for the link of perinatal distress to pregnancy outcomes (i.e., still birth, spontaneous abortion, miscarriage, and preterm birth) and infant health. Perinatal distress was defined as the psychological response to episodic (e.g., traumatic life events, catastrophic events) stress, chronic stress (e.g., adverse socioeconomic, cultural, and environmental phenomena), and pregnancy-related mood disorders (e.g., depression, anxiety). The author suggested that all dimensions of perinatal distress should be captured to examine its impact on pregnancy outcomes. Olson et al. (2015) constructed a conceptual framework to link transgenerational AL and preterm birth. In their model, generational experience is acquired via genetic and epigenetic inheritance. Cumulative life stressors (i.e., early life adversity, social context, stress exposure) over the life course cause behavioral, psychological, and inflammatory stress responses leading to high AL, and ultimately increase the risk for preterm birth.

The two models by Premji (2014) and Olson et al. (2015) applied the AL theory in the childbearing context to understand the etiologic contribution of perinatal distress or cumulative life stressors on birth outcomes. However, they did not consider how the three stages of the AL process (i.e., primary mediators, secondary outcomes, tertiary outcomes) could be specified for pregnancy. In addition, it has been suggested to distinguish the two levels of mediation (primary mediators and secondary outcomes) in AL research (Juster, McEwen, & Lupien, 2010). Primary mediators are more related to the stressors and reflect the short-term effects of stress (i.e., cortisol release from neuroendocrine system), while secondary outcomes represent the long-term effects of stress and are signs of dysregulation (i.e., abnormal levels of biomarkers in

cardiovascular, metabolic, and immune systems). Therefore, our newly synthesized model of perinatal AL now builds on the 4 aforementioned models by encompassing the different stages of the AL process. Our proposed model focuses on the impact of maternal pre-existing and current PTSD—a particularly severe, chronic, and treatable form of stress—and considers the influence of other maternal risk factors on adverse birth outcomes associated with maternal PTSD. Table 2-1 provides brief descriptions of the four approaches to stress-outcome.

Empirical basis for main propositions

Maternal PTSD, AL, pregnancy complications, and adverse birth outcomes

In a national sample of U.S. adults, lifetime, past-12 month, and past-6 month PTSD prevalences were much higher among women than men (Kilpatrick et al., 2013). The prevalence of PTSD was particularly higher in African American pregnant women than their non-African American pregnant counterparts (Seng, Kohn-Wood, McPherson, & Sperlich, 2011a). Although previous findings were inconsistent due to some limitations, three relatively recent studies with large representative samples consistently found PTSD as a severe, chronic form of disordered stress regulation might be associated with increased risk of adverse birth outcomes (Seng et al., 2011b; Shaw et al., 2014; Yonkers et al., 2014). The recent study by Shaw et al. (2017) also found women with current PTSD diagnosis had elevated risk of pregnancy complications (e.g., preeclampsia, gestational diabetes). Preeclampsia has been reported as the primary maternal cause of medically indicated preterm birth (Ananth and Vintzileos, 2006) and gestational diabetes as a risk factor of preeclampsia were also found in relation to increased risk of medically induced premature delivery (Institute of Medicine, 2007).

Applying AL theory to the childbearing context, we posit that maternal pre-existing and current PTSD as a chronic stress could initiate the activities of multiple physiological systems.

Dysregulated production of primary mediators gives rise to secondary outcomes including changes in immune, metabolic, and cardiovascular systems, and anthropometric parameters. The overall accumulated physiological dysregulation could either directly result in quaternary outcomes (e.g., preterm birth and low birth weight) or indirectly lead to pregnancy-specific tertiary outcomes (e.g., preeclampsia, gestational diabetes, infection, and premature rupture of membrane) that ultimately cause quaternary outcomes. Overall, it may be theoretically posited that AL mediates the relationships between maternal PTSD and birth outcomes (Figure 2-2). As a biologically mediated pathway between stress and health, AL may be a possible contributor to adverse birth outcomes in women with PTSD.

To our knowledge, there are only three studies exploring the link between AL and adverse birth outcomes, but the results were inconsistent. Findings from the Bogalusa Heart Study suggested that AL was not associated with preterm birth or low birth weight in black and white women (Wallace et al., 2013b). On the other hand, a case-control study using the National Health and Nutrition Examination Survey (NHANES) data demonstrated that non-pregnant women with history of small for gestational age or preterm birth had higher AL levels compared to those with normal birth outcomes (Hux, Catov, & Roberts, 2014). Another study by Wallace et al. (2013a) conducted in New Orleans found negative associations of AL with gestational age among 42 women. Each of these three studies had methodological limitations that may explain the incongruent findings: the AL data was collected before pregnancy (Wallace et al., 2013b) or during non-pregnancy (Hux et al., 2014), or the sample size was small (Wallace et al., 2013a). Except for the three studies, a study using data from the Prenatal Exposures and Preeclampsia Prevention (PEPP) study found women with higher AL had increasing odds of developing preeclampsia (Hux and Roberts, 2015). In summary, there are inconsistencies with study results

with one study finding non-significant associations between preconception AL and preterm birth and low birth weight; another finding significant associations, but AL data was not collected during pregnancy; and the other two studies reporting significant associations between pregnancy AL and gestational age and preeclampsia.

Although only a few studies tested AL levels in relation to birth outcomes, AL shows promise as a theoretical framework for predicting birth outcomes (Olson et al., 2015; Premji, 2014; Shannon, King, & Kennedy, 2007). Elevated cortisol levels as the primary indicator of AL have been well documented in women with adverse birth outcomes (Braig et al., 2015; Goedhart et al., 2010). Premji (2014) reviewed studies that examined the relationship between prenatal stress, biomarkers of stress, and preterm birth and found individual biomarkers from neuroendocrine, immune, and cardiovascular systems were associated with preterm birth. Olson et al. (2015) also reviewed animal studies that demonstrated the impact of stress, stress hormones, and inflammatory mediators on preterm birth.

No studies have examined AL in relation to PTSD among pregnant women. But one study with mothers of pediatric cancer survivors as a high-stress population indicated elevated AL levels in women with high-stress histories, particularly those with PTSD (Glover, Stuber, & Poland, 2006), suggesting there is a relationship between PTSD and AL in a female population. However, no studies have investigated AL as a pathway for the impact of maternal PTSD on birth outcomes. Thus, future studies are needed to test whether AL can serve as a potential contributor to adverse birth outcomes in women with PTSD. Building a theoretical model is the first step in establishing this line of research.

Child maltreatment, IPV, SES, depression, pregnancy-related distress, and risk behaviors

Some maternal risk factors including child maltreatment, intimate partner violence (IPV), low SES, depression, pregnancy-related distress, and risk behaviors may influence the impact of maternal PTSD on birth outcomes. Child maltreatment is a common trauma occurring in early life. It has been found to pose the greatest risk for PTSD during pregnancy (Seng, Low, Sperlich, Ronis, & Liberzon, 2009). The associations between PTSD and adverse birth outcomes were stronger in pregnant women with histories of child maltreatment than other types of trauma (Seng et al., 2011b). IPV against women has been recognized as a global problem due to its detrimental effect on women and their children (Koenig, Stephenson, Ahmed, Jejeebhoy, & Campbell, 2006). IPV as an adulthood trauma is a strong predictor of PTSD in women (Gobin, Iverson, Mitchell, Vaughn, & Resick, 2013). Sociodemographic disadvantage including low SES has great influence on individual's health (Kawada, 2014; Leng, Jin, Li, Chen, & Jin, 2014). DiGrande et al. (2008) found that lower SES could increase risk for PTSD among residents highly exposed to the 9/11 terrorist attacks. Both IPV exposure and lower SES were found in relation to adverse birth outcomes (Smith, Draper, Manktelow, Dorling, & Field, 2007; Watson and Taft, 2013). Thus, child maltreatment, IPV, and low SES as chronic stress may increase the risk of maternal PTSD that eventually contributes to adverse birth outcomes.

PTSD is frequently comorbid with depression and risk behaviors (e.g., tobacco, and drug use; Creamer, Burgess, & McFarlane, 2001; Kessler, Chiu, Demler, Merikangas, & Walters, 2005b; Kessler, Sonnega, Bromet, Hughes, & Nelson, 1995; Lopez & Seng, 2014; Yildiz et al., 2017). Pregnancy-related distress such as anxiety refers to women's fear about unintended pregnancy, their infants' health, their own health, and delivery. Depression, risk behaviors, and pregnancy-related distress have been identified as maternal risk factors for poor birth outcomes

(Andres & Day, 2000; Cnattingius, 2004; Goldenberg et al., 2008; Orr & Miller, 1995; Schetter & Tanner, 2012). But the potential interactions between those maternal risk factors and PTSD on adverse birth outcomes are unknown. Given the high comorbidity of PTSD and those risk factors, they may serve as moderators to strengthen the negative effect of maternal PTSD on birth outcomes.

Furthermore, some literature reviews also support the relationships between child maltreatment, IPV, low SES, depression, and risk behaviors and higher levels of AL (Danese & McEwen, 2012; Dowd, Simanek, & Aiello, 2009; McEwen, 2003; McEwen & Tucker, 2011; Weiss, 2007). Taken together, it can be hypothesized that child maltreatment, IPV, and low SES as chronic stress could increase the risk of developing maternal PTSD. Depression, risk behaviors, and pregnancy-related distress as moderators could enhance the impact of maternal PTSD on AL that eventually contribute to poor birth outcomes (Figure 2-3). Including those maternal risk factors in the theoretical model could add our understanding of the associations between maternal PTSD and adverse birth outcomes.

Explication of a new perinatal AL theoretical model

The perinatal AL theoretical model was constructed by integrating elements of the 4 existing theoretical models (Beckie, 2012; Seng, 2002; Premji, 2014; Olson et al., 2015) and limited empirical evidence. The AL theoretical framework by Beckie (2012) depicts that allostatic challenges could lead to tertiary outcomes through the accumulated multisystem physiological dysfunctions including changes in primary mediators and secondary outcomes. The focal concepts in the theoretical framework include allostatic challenges, primary mediators, secondary outcomes, and tertiary outcomes. Applying the theoretical framework in the childbearing context, we included not only the same concepts but also added quaternary

outcomes as a new concept in our model based upon empirical evidence. Therefore, the focal concepts within our synthesized model are allostatic challenges, primary mediators, secondary outcomes, tertiary outcomes, and quaternary outcomes.

The variables related to those focal concepts were included in our model based on the 4 previous models and empirical evidence. Consistent with the conceptual framework by Seng (2002), our model focuses on maternal pre-existing and current PTSD as a primary allostatic challenge. We also included child maltreatment, IPV, SES, maternal depression, pregnancy-related distress, and risk behaviors (i.e., tobacco, alcohol, and drug use) as other allostatic challenges in our model, which is congruent with the previous models (Beckie, 2012; Premji, 2014; Olson et al., 2015). Primary mediators are physiological indicators from neuroendocrine system such as cortisol, epinephrine, and norepinephrine. Secondary outcomes are physiological indicators from immune, metabolic, and cardiovascular systems (e.g., c-reactive protein, glucose, blood pressure) and anthropometric parameters (e.g., body mass index, waist/hip ratio). The tertiary outcomes are pregnancy-specific, including preeclampsia, gestational diabetes, infections, and premature rupture of membrane, and the quaternary outcomes are adverse birth outcomes (i.e., preterm birth and low birth weight).

Child maltreatment, IPV, low SES are antecedents of PTSD; maternal depression, pregnancy-related distress, and risk behaviors serve as moderators for the impact of maternal PTSD on birth outcomes; and primary mediators, secondary outcomes, and tertiary outcomes play a mediating role in the impact of maternal PTSD on birth outcomes. Because the proposed model focused on the impact of maternal PTSD and its pathways on adverse birth outcomes, the direct links from maternal depression, pregnancy-related distress, and risk behaviors to adverse birth outcomes were not taken into account. The depiction of the proposed new perinatal AL

model was as follows: Child maltreatment, IPV, low SES as chronic stress could increase the risk for developing PTSD. As a result of maternal PTSD, the dysregulated production of primary mediators (e.g., cortisol, epinephrine, norepinephrine) leads to secondary outcomes including dysregulations in immune, metabolic, and cardiovascular systems, and changes in anthropometric parameters. The accumulated multisystem physiological dysregulations could either directly cause quaternary outcomes of poor birth outcomes (i.e., preterm birth and low birth weight) or indirectly result in tertiary outcomes (i.e., preeclampsia, gestational diabetes, infection, premature rupture of membrane) that eventually lead to quaternary outcomes. Depression comorbidity, pregnancy-related distress, and risk behaviors as moderators could strengthen the impact of maternal PTSD on birth outcomes through the AL process. According to the model by Olson et al. (2015), intergenerational patterns (i.e., genetics and epigenetics) also play a significant role in the pathway from maternal PTSD to adverse birth outcomes. The synthesized model is depicted in Table 2-2 and Figure 2-4.

Evaluation of the new perinatal AL theoretical model

This new perinatal AL theoretical model consists of concepts and variables across psychosocial, behavioral, psychopathological, and neurobiological perspectives. It is challenging to build bridges between different sciences. The complexity of neurobiological systems makes it even more challenging to understand the nature of the relationships. The proposed new model presented the links from maternal stress to birth outcomes through AL, but the inter-relationships among concepts and variables may be more complex. Multiple physiological indicators from different body systems might be interrelated in a non-linear network and impose non-linear effects on health (Misra, Straughen, Slaughter-Acey, 2013). In our model, maternal stress/PTSD causes complications in labor/pregnancy, but in turn perinatal complications as a type of trauma

may increase the risk for developing PTSD (Andersen, Melvaer, Videbech, Lamont, Joergensen, 2012). Thus, much work is still needed to better understand the inter-relationships among different concepts. The proposed new model needs to be validated and continuously modified based on empirical evidence. The contribution of our model is to provide an explanatory framework and a direction for future research to further explore the potential mechanisms for the impact of maternal PTSD on adverse birth outcomes.

Implications for research

The theoretical model synthesized in this paper is useful for its potential to advance perinatal research. It will enable researchers to concurrently examine maternal stress and physiological responses that influence birth outcomes. Based on the theoretical model, prospective studies with multiple biological and psychosocial measures of PTSD and maternal risk factors (i.e., child maltreatment, IPV, lower SES, depression, pregnancy-related distress, risk behaviors) may add new knowledge about the etiologic contributions of psychosocial processes to adverse birth outcomes. Several steps need to be taken to advance application of the theoretical model in perinatal research. First, the selection of physiological indicators that have the most predictive performance on birth outcomes and the timing of data collection during pregnancy need to be determined. Many combinations of biomarkers and anthropometric parameters have been used to create the AL score. There is no validated gold standard measure of AL, and this is especially true in pregnancy. Selection of AL indicators should be determined with careful consideration for pregnancy physiology to avoid the inclusion of indicators that are not relevant to physiological dysregulation or exclusion relevant measures. It is essential to develop a panel of optimal and valid AL indicators that can be included as part of routine prenatal care in the future. Due to the changing levels of physiological measures across

pregnancy, it is also unknown the optimal timing of data collection. Multiple assessments across pregnancy take a lot of time, efforts, and financial resources and may lead to lots of missing data as well. The challenges increase when dealing with high risk and vulnerable pregnant women. It would be feasible if we could determine the optimal timing of data collection at which largest individual differences (variances) can be tested. Therefore, future studies are needed to determine the best combination of AL indicators and the optimal timing of specimen collection. Second, the optimal AL scoring method needs to be determined. Despite continued work using various scoring methods to construct the AL index, few studies focused on validating the AL index and comparing which method is optimal to capture the multiple and inter-connected feature of AL and summarize such information into one cumulative index. Such comparative data remains sparse and the question of how best to score AL remains to be addressed. The lack of comparative analyses of the reliability of different AL measures among pregnant women represents a gap in the current research on maternal stress and birth outcomes. Thus, more comparative work is clearly needed to compare the explanatory powers and predictive abilities of distinct scoring approaches and to determine the optimal scoring approach before the theoretical model can be reliably and validly applied to perinatal outcome research. Lastly, after these methodological issues are addressed, prospective studies with multiple biological and psychosocial measures of stress are needed to test the theoretical model. The proposed model needs to be assessed, potentially modified, and validated to sort out all the relationships among concepts and variables.

Potential implications for clinical practice, service delivery, and policy

After the synthesized model is well validated, it may have potential to influence attention to clinical practice, service delivery, and policy. Women with PTSD as well as those exposed to

the co-occurring risk factors have been a neglected population in studies on perinatal distress and pregnancy and infant outcomes (Premji, 2014). PTSD affects women in low resource settings at higher rates (Seng et al., 2009), and these are women with less limited access to health care. Applying the theoretical model to clinical practice and service delivery would help health care providers including obstetricians and midwives be aware of PTSD as a risk factor of adverse birth outcomes, perceive the associations between stress, PTSD, health alterations that accumulate and interact, and adverse birth outcomes, and pay more attention to pregnant women with PTSD. Brief, effective screening tools for PTSD (Ouimette, Wade, Prins, & Schohn, 2008) could be included in prenatal care, especially among high-risk women (i.e., those exposed to child maltreatment, IPV, lower SES, depression, pregnancy-related distress, risk behaviors). Appropriate mental health care and education about how to identify and reduce stress and traumatic stress reactions could also be provided for high-risk women in the prenatal period in order to address stress dysregulation and improve both maternal well being and infant outcomes. Since mental health care services may be nonexistent or limited and predominantly hospital based (Saxena et al., 2011), consideration could be given to developing or strengthening local mental health care services that will serve women during the perinatal period. Finally, policies can be worked out to mandate screening for PTSD and providing mental health service in pregnant women. Emphasizing health care for the vulnerable population in public policy would help this vulnerable population seek and receive adequate care to prevent adverse birth outcomes.

Conclusion

The proposed theoretical model could provide a direction for researchers to further explore the biological mechanisms by which maternal chronic stress such as PTSD leads to

adverse birth outcomes. Exploring the contribution of maternal chronic stress to adverse birth outcomes within the context of a theoretical model may provide evidence for health care professionals to identify individuals at risks for adverse birth outcomes, and thus determine what diagnosis, prevention, or treatment are good for pregnant women to prevent those adverse outcomes.

References

- Allen, M. C. (2008). Neurodevelopmental outcomes of preterm infants. *Current Opinion in Neurology*, 21(2), 123-128. doi: 10.1097/WCO.0b013e3282f88bb4
- Andersen, L. B., Melvaer, L. B., Videbech, P., Lamont, R. F., Joergensen, J. S. (2012). Risk factors for developing post-traumatic stress disorder following childbirth: a systematic review. *Acta obstetricia et gynecologica Scandinavica*, 91(11), 1261-1272.
- Ananth, C.V., Vintzileos, A.M. (2006). Maternal-fetal conditions necessitating a medical intervention resulting in preterm birth. *American Journal of Obstetrics & Gynecology* 195, 1557–1563.
- Andres, R. L., & Day, M. C. (2000). Perinatal complications associated with maternal tobacco use. *Seminars in Neonatology*, 5(3), 231-241. doi: 10.1053/siny.2000.0025
- Beckie, T. M. (2012). A systematic review of allostatic load, health, and health disparities. *Biological Research for Nursing*, 14(4), 311-346.
- Berkowitz, G. S., Wolff, M. S., Janevic, T. M., Holzman, I. R., Yehuda, R., & Landrigan, P. J. (2003). The World Trade Center disaster and intrauterine growth restriction. *JAMA*, 290(5), 595-596. doi: 10.1001/jama.290.5.595-b
- Borrell, L. N., Dallo, F. J., & Nguyen, N. (2010). Racial/ethnic disparities in all-cause mortality in U.S. adults: the effect of allostatic load. *Public Health Reports*, 125(6), 810-816.
- Braig, S., Grabher, F., Ntomchukwu, C., Reister, F., Stalder, T., Kirschbaum, C., . . . Rothenbacher, D. (2015). Determinants of maternal hair cortisol concentrations at delivery reflecting the last trimester of pregnancy. *Psychoneuroendocrinology*, 52, 289-296. doi: 10.1016/j.psyneuen.2014.12.006
- Chang, H. L., Chang, T. C., Lin, T. Y., & Kuo, S. S. (2002). Psychiatric morbidity and pregnancy outcome in a disaster area of Taiwan 921 earthquake. *Psychiatry and Clinical Neurosciences*, 56(2), 139-144. doi: 10.1046/j.1440-1819.2002.00948.x
- Chyu, L., & Upchurch, D. M. (2011). Racial and ethnic patterns of allostatic load among adult women in the United States: findings from the National Health and Nutrition Examination Survey 1999-2004. *Journal of Women's Health (Larchmt)*, 20, 575-583.
- Cnattingius, S. (2004). The epidemiology of smoking during pregnancy: smoking prevalence, maternal characteristics, and pregnancy outcomes. *Nicotine & Tobacco Research*, 6 Suppl 2, S125-140. doi: 10.1080/14622200410001669187
- Creamer, M., Burgess, P., & McFarlane, A. C. (2001). Post-traumatic stress disorder: findings from the Australian National Survey of Mental Health and Well-being. *Psychological Medicine*, 31(7), 1237-1247.
- Crimmins, E. M., Kim, J. K., & Seeman, T. E. (2009). Poverty and Biological Risk: The Earlier "Aging" of the Poor. *Journals of Gerontology Series a-Biological Sciences and Medical Sciences*, 64(2), 286-292. doi: 10.1093/gerona/gln010
- Danese, A., & McEwen, B. S. (2012). Adverse childhood experiences, allostasis, allostatic load, and age-related disease. *Physiology & Behavior*, 106(1), 29-39. doi: 10.1016/j.physbeh.2011.08.019

- DiGrande, L., Perrin, M. A., Thorpe, L. E., Thalji, L., Murphy, J., Wu, D., . . . Brackbill, R. M. (2008). Posttraumatic stress symptoms, PTSD, and risk factors among lower Manhattan residents 2-3 years after the September 11, 2001 terrorist attacks. *Journal of Traumatic Stress, 21*(3), 264-273. doi: 10.1002/jts.20345
- Dowd, J. B., Simanek, A. M., & Aiello, A. E. (2009). Socio-economic status, cortisol and allostatic load: a review of the literature. *International Journal of Epidemiology, 38*(5), 1297-1309. doi: 10.1093/ije/dyp277
- Engel, S. M., Berkowitz, G. S., Wolff, M. S., & Yehuda, R. (2005). Psychological trauma associated with the World Trade Center attacks and its effect on pregnancy outcome. *Paediatric and Perinatal Epidemiology, 19*(5), 334-341. doi: 10.1111/j.1365-3016.2005.00676.x
- Glover, D. A., Stuber, M., & Poland, R. E. (2006). Allostatic load in women with and without PTSD symptoms. *Psychiatry, 69*(3), 191-203.
- Gobin, R. L., Iverson, K. M., Mitchell, K., Vaughn, R., & Resick, P. A. (2013). The impact of childhood maltreatment on PTSD symptoms among female survivors of intimate partner violence. *Violence and Victims, 28*(6), 984-999.
- Goedhart, G., Vrijkotte, T. G., Roseboom, T. J., van der Wal, M. F., Cuijpers, P., & Bonsel, G. J. (2010). Maternal cortisol and offspring birthweight: results from a large prospective cohort study. *Psychoneuroendocrinology, 35*(5), 644-652. doi: 10.1016/j.psyneuen.2009.10.003
- Goldenberg, R. L., Culhane, J. F., Iams, J. D., & Romero, R. (2008). Epidemiology and causes of preterm birth. *Lancet, 371*(9606), 75-84. doi: 10.1016/S0140-6736(08)60074-4
- Goldman, N., Turra, C. M., Gleib, D. A., Lin, Y. H., & Weinstein, M. (2006). Physiological dysregulation and changes in health in an older population. *Experimental Gerontology, 41*(9), 862-870. doi: 10.1016/j.exger.2006.06.050
- Green, N. S., Damus, K., Simpson, J. L., Iams, J., Reece, E. A., Hobel, C. J., . . . March Of Dimes Scientific Advisory Committee On, P. (2005). Research agenda for preterm birth: recommendations from the March of Dimes. *American Journal of Obstetrics & Gynecology, 193*(3 Pt 1), 626-635. doi: 10.1016/j.ajog.2005.02.106
- Gruenewald, T. L., Seeman, T. E., Karlamangla, A. S., & Sarkisian, C. A. (2009). Allostatic load and frailty in older adults. *Journal of the American Geriatrics Society, 57*(9), 1525-1531. doi: <http://dx.doi.org/10.1111/j.1532-5415.2009.02389.x>
- Gustafsson, P. E., Janlert, U., Theorell, T., Westerlund, H., & Hammarstrom, A. (2011). Socioeconomic status over the life course and allostatic load in adulthood: results from the Northern Swedish Cohort. *Journal of Epidemiology & Community Health, 65*(11), 986-992. doi: <http://dx.doi.org/10.1136/jech.2010.108332>
- Gustafsson, P. E., Janlert, U., Theorell, T., Westerlund, H., & Hammarstrom, A. (2012). Social and material adversity from adolescence to adulthood and allostatic load in middle-aged women and men: results from the Northern Swedish Cohort. *Annals of Behavioral Medicine, 43*(1), 117-128. doi: 10.1007/s12160-011-9309-6
- Hellgren, C., Akerud, H., Skalkidou, A., & Sundstrom-Poromaa, I. (2013). Cortisol awakening response in late pregnancy in women with previous or ongoing depression. *Psychoneuroendocrinology, 38*(12), 3150-3154. doi: 10.1016/j.psyneuen.2013.08.007
- Hux, V. J., Catov, J. M., & Roberts, J. M. (2014). Allostatic load in women with a history of low birth weight infants: the national health and nutrition examination survey. *Journal of Women's Health (Larchmt), 23*(12), 1039-1045. doi: 10.1089/jwh.2013.4572

- Hux, V. J., & Roberts, J. M. (2015). A potential role for allostatic load in preeclampsia. *Maternal & Child Health Journal, 19*(3), 591-597. doi: 10.1007/s10995-014-1543-7
- Institute of Medicine (2007). *Preterm Birth: causes, Consequences, and Prevention*. National Academies Press, Washington, D.C. pp. 148–168.
- Juster, R. P., McEwen, B. S., & Lupien, S. J. (2010). Allostatic load biomarkers of chronic stress and impact on health and cognition. *Neuroscience & Biobehavioral Reviews, 35*(1), 2-16. doi: <http://dx.doi.org/10.1016/j.neubiorev.2009.10.002>
- Kawada, T. (2014). Socioeconomic status and obesity: Causality of the association. *Journal of Epidemiologic Global Health, 4*(4), 327-328. doi: 10.1016/j.jegh.2014.07.005
- Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005a). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62*(6), 593-602. doi: 10.1001/archpsyc.62.6.593
- Kessler, R. C., Chiu, W. T., Demler, O., Merikangas, K. R., & Walters, E. E. (2005b). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62*(6), 617-627. doi: 10.1001/archpsyc.62.6.617
- Kessler, R. C., Sonnega, A., Bromet, E., Hughes, M., & Nelson, C. B. (1995). Posttraumatic stress disorder in the National Comorbidity Survey. *Archives of General Psychiatry, 52*(12), 1048-1060.
- Kilpatrick, D. G., Resnick, H. S., Milanak, M. E., Miller, M. W., Keyes, K. M., & Friedman, M. J. (2013). National Estimates of Exposure to Traumatic Events and PTSD Prevalence Using DSM-IV and DSM-5 Criteria. *Journal of Traumatic Stress, 26*(5), 537-547. doi: 10.1002/jts.21848
- King, N. M., Chambers, J., O'Donnell, K., Jayaweera, S. R., Williamson, C., & Glover, V. A. (2010). Anxiety, depression and saliva cortisol in women with a medical disorder during pregnancy. *Archives of Women's Mental Health, 13*(4), 339-345. doi: 10.1007/s00737-009-0139-5
- Koenig, M. A., Stephenson, R., Ahmed, S., Jejeebhoy, S. J., & Campbell, J. (2006). Individual and contextual determinants of domestic violence in North India. *American Journal of Public Health, 96*(1), 132-138. doi: 10.2105/AJPH.2004.050872
- Latendresse, G. (2009). The interaction between chronic stress and pregnancy: preterm birth from a biobehavioral perspective. *Journal of Midwifery & Women's Health, 54*(1), 8-17. doi: <http://dx.doi.org/10.1016/j.jmwh.2008.08.001>
- Lawn, J. E., Gravett, M. G., Nunes, T. M., Rubens, C. E., Stanton, C., & Group, G. R. (2010). Global report on preterm birth and stillbirth (1 of 7): definitions, description of the burden and opportunities to improve data. *BMC Pregnancy Childbirth, 10 Suppl 1*, S1. doi: 10.1186/1471-2393-10-S1-S1
- Lawn, J. E., Wilczynska-Ketende, K., & Cousens, S. N. (2006). Estimating the causes of 4 million neonatal deaths in the year 2000. *International Journal of Epidemiology, 35*(3), 706-718. doi: 10.1093/ije/dyl043
- Leng, B., Jin, Y., Li, G., Chen, L., & Jin, N. (2014). Socioeconomic status and hypertension: a meta-analysis. *Journal of Hypertension*. doi: 10.1097/HJH.0000000000000428
- Lipkind, H. S., Curry Ae Fau - Huynh, M., Huynh M Fau - Thorpe, L. E., Thorpe Le Fau - Matte, T., & Matte, T. (2010). Birth outcomes among offspring of women exposed to the September 11, 2001, terrorist attacks. *Obstetrics & Gynecology, 116*(4), 917-925.

- Lopez, W. D., & Seng, J. S. (2014). Posttraumatic stress disorder, smoking, and cortisol in a community sample of pregnant women. *Addictive Behavior, 39*(10), 1408-1413. doi: 10.1016/j.addbeh.2014.04.027
- Martin, J. A., Hamilton, B. E., Osterman, M. J., Curtin, S. C., & Matthews, T. J. (2015). Births: final data for 2013. *National Vital Statistics Reports, 64*(1), 1-65.
- Martin, J. A., Hamilton, B. E., Ventura, S. J., Osterman, M. J., Wilson, E. C., & Mathews, T. J. (2012). Births: final data for 2010. *National Vital Statistics Reports, 61*(1), 1-72.
- McEwen, B. S. (1998). Protective and damaging effects of stress mediators. *The New England Journal of Medicine, 338*(3), 171-179. doi: 10.1056/NEJM199801153380307
- McEwen, B. S. (2000). Allostasis and allostatic load: implications for neuropsychopharmacology. *Neuropsychopharmacology, 22*(2), 108-124.
- McEwen, B. S. (2003). Mood disorders and allostatic load. *Biological Psychiatry, 54*(3), 200-207.
- McEwen, B. S. (2006). Protective and damaging effects of stress mediators: central role of the brain. *Dialogues in Clinical Neuroscience, 8*(4), 367-381.
- McEwen, B. S., & Seeman, T. (1999). Protective and damaging effects of mediators of stress: elaborating and testing the concepts of allostasis and allostatic load. *Annals of the New York Academy of Sciences, 896*(1), 30-47.
- McEwen, B. S., & Stellar, E. (1993). Stress and the individual. Mechanisms leading to disease. *Archives of Internal Medicine, 153*(18), 2093-2101.
- McEwen, B. S., & Tucker, P. (2011). Critical biological pathways for chronic psychosocial stress and research opportunities to advance the consideration of stress in chemical risk assessment. *American Journal of Public Health, 101 Suppl 1*, S131-139. doi: 10.2105/AJPH.2011.300270
- McEwen, B. S., & Wingfield, J. C. (2003). The concept of allostasis in biology and biomedicine. *Hormones & Behavior, 43*(1), 2-15.
- Misra, D. P., Straughen, J. K., Slaughter-Acey, J. C. (2013). Allostatic load and health: Can perinatal epidemiology lead the way forward?. *Paediatric and perinatal epidemiology, 27*(6), 507-508.
- Morland, L., Goebert, D., Onoye, J., Frattarelli, L., Derauf, C., Herbst, M., . . . Friedman, M. (2007). Posttraumatic stress disorder and pregnancy health: preliminary update and implications. *Psychosomatics, 48*(4), 304-308. doi: 10.1176/appi.psy.48.4.304
- Olson, D. M., Severson, E. M., Verstraeten, B. S., Ng, J. W., McCreary, J. K., & Metz, G. A. (2015). Allostatic Load and Preterm Birth. *International Journal of Molecular Science, 16*(12), 29856-29874. doi: 10.3390/ijms161226209
- Orr, S. T., & Miller, C. A. (1995). Maternal depressive symptoms and the risk of poor pregnancy outcome. Review of the literature and preliminary findings. *Epidemiologic Review, 17*(1), 165-171.
- Ouimette, P., Wade, M., Prins, A., & Schohn, M. (2008). Identifying PTSD in primary care: comparison of the Primary Care-PTSD screen (PC-PTSD) and the General Health Questionnaire-12 (GHQ). *Journal of Anxiety Disorder, 22*(2), 337-343. doi: 10.1016/j.janxdis.2007.02.010
- Premji, S. (2014). Perinatal distress in women in low- and middle-income countries: allostatic load as a framework to examine the effect of perinatal distress on preterm birth and infant health. *Maternal & Child Health Journal, 18*(10), 2393-2407. doi: 10.1007/s10995-014-1479-y

- Resnick, H. S., Kilpatrick, D. G., Dansky, B. S., Saunders, B. E., & Best, C. L. (1993). Prevalence of civilian trauma and posttraumatic stress disorder in a representative national sample of women. *Journal of Consulting and Clinical Psychology, 61*(6), 984-991.
- Risnes, K. R., Vatten, L. J., Baker, J. L., Jameson, K., Sovio, U., Kajantie, E., . . . Bracken, M. B. (2011). Birthweight and mortality in adulthood: a systematic review and meta-analysis. *International Journal of Epidemiology, 40*(3), 647-661. doi: 10.1093/ije/dyq267
- Rogal, S. S., Poschman, K., Belanger, K., Howell, H. B., Smith, M. V., Medina, J., & Yonkers, K. A. (2007). Effects of posttraumatic stress disorder on pregnancy outcomes. *Journal of Affective Disorders, 102*(1-3), 137-143. doi: 10.1016/j.jad.2007.01.003
- Rosemberg, M. A. S., Li, Y., & Seng, J., 2017. Allostatic load: a useful concept for advancing nursing research. *Journal of clinical nursing 26*(23-24), 5191-5205.
- Rosen, D., Seng Js Fau - Tolman, R. M., Tolman Rm Fau - Mallinger, G., & Mallinger, G. (2007). Intimate partner violence, depression, and posttraumatic stress disorder as additional predictors of low birth weight infants among low-income mothers. *Journal of Interpersonal Violence, 22*(10), 1305-1314.
- Saxena, S., Lora, A., Morris, J., Berrino, A., Esparza, P., Barrett, T., . . . Saraceno, B. (2011). Mental health services in 42 low- and middle-income countries: a WHO-AIMS cross-national analysis. *Psychiatric Service, 62*(2), 123-125. doi: 10.1176/appi.ps.62.2.123
- Schetter, C. D., & Tanner, L. (2012). Anxiety, depression and stress in pregnancy: implications for mothers, children, research, and practice. *Current Opinion in Psychiatry, 25*(2), 141-148. doi: 10.1097/YCO.0b013e3283503680
- Seeman, T., Merkin, S. S., Crimmins, E., Koretz, B., Charette, S., & Karlamangla, A. (2008). Education, income and ethnic differences in cumulative biological risk profiles in a national sample of US adults: NHANES III (1988-1994). *Social Science & Medicine, 66*(1), 72-87. doi: 10.1016/j.socscimed.2007.08.027
- Seeman, T. E., Crimmins, E., Huang, M. H., Singer, B., Bucur, A., Gruenewald, T., . . . Reuben, D. B. (2004). Cumulative biological risk and socio-economic differences in mortality: MacArthur studies of successful aging. *Social Science & Medicine, 58*(10), 1985-1997.
- Seeman, T. E., McEwen, B. S., Rowe, J. W., & Singer, B. H. (2001). Allostatic load as a marker of cumulative biological risk: MacArthur studies of successful aging. *Proceedings of the National Academy of Sciences of the United States of America, 98*(8), 4770-4775.
- Seeman, T. E., Singer, B. H., Rowe, J. W., Horwitz, R. I., & McEwen, B. S. (1997). Price of adaptation--allostatic load and its health consequences. MacArthur studies of successful aging. *Archives of Internal Medicine, 157*(19), 2259-2268.
- Seng, J. S. (2002). A conceptual framework for research on lifetime violence, posttraumatic stress, and childbearing. *Journal of Midwifery & Women's Health, 47*(5), 337-346.
- Seng, J. S., Kohn-Wood, L. P., McPherson, M. D., & Sperlich, M. (2011a). Disparity in posttraumatic stress disorder diagnosis among African American pregnant women. *Archives of Women's Mental Health, 14*(4), 295-306. doi: 10.1007/s00737-011-0218-2
- Seng, J. S., Low, L. K., Sperlich, M., Ronis, D. L., & Liberzon, I. (2009). Prevalence, trauma history, and risk for posttraumatic stress disorder among nulliparous women in maternity care. *Obstetrics & Gynecology, 114*(4), 839-847. doi: 10.1097/AOG.0b013e3181b8f8a2
- Seng, J. S., Low, L. K., Sperlich, M., Ronis, D. L., & Liberzon, I. (2011b). Post-traumatic stress disorder, child abuse history, birthweight and gestational age: a prospective cohort study.

- BJOG: An International Journal of Obstetrics & Gynaecology*, 118(11), 1329-1339. doi: 10.1111/j.1471-0528.2011.03071.x
- Seng, J. S., Oakley, D. J., Sampselle, C. M., Killion, C., Graham-Bermann, S., & Liberzon, I. (2001). Posttraumatic stress disorder and pregnancy complications. *Obstetrics & Gynecology*, 97(1), 17-22.
- Seplaki, C. L., Goldman, N., Weinstein, M., & Lin, Y. H. (2006). Measurement of cumulative physiological dysregulation in an older population. *Demography*, 43(1), 165-183.
- Shannon, M., King, T. L., & Kennedy, H. P. (2007). Allostasis: a theoretical framework for understanding and evaluating perinatal health outcomes. *JOGNN - Journal of Obstetric, Gynecologic, & Neonatal Nursing*, 36(2), 125-134.
- Shaw, J. G., Asch, S. M., Kimerling, R., Frayne, S. M., Shaw, K. A., & Phibbs, C. S. (2014). Posttraumatic stress disorder and risk of spontaneous preterm birth. *Obstetrics & Gynecology*, 124(6), 1111-1119. doi: 10.1097/AOG.0000000000000542
- Shaw, J. G., Asch, S. M., Katon, J. G., Shaw, K. A., Kimerling, R., Frayne, S. M., Phibbs, C. S. (2017). Post-traumatic Stress Disorder and Antepartum Complications: a Novel Risk Factor for Gestational Diabetes and Preeclampsia. *Paediatric Perinatal Epidemiology*, 31(3), 185-194. doi: 10.1111/ppe.12349
- Shea, A. K., Streiner, D. L., Fleming, A., Kamath, M. V., Broad, K., & Steiner, M. (2007). The effect of depression, anxiety and early life trauma on the cortisol awakening response during pregnancy: Preliminary results. *Psychoneuroendocrinology*, 32(8-10), 1013-1020. doi: 10.1016/j.psyneuen.2007.07.006
- Shonkoff, J. P., Garner, A. S., Committee on Psychosocial Aspects of, C., Family, H., Committee on Early Childhood, A., Dependent, C., . . . Behavioral, P. (2012). The lifelong effects of early childhood adversity and toxic stress. *Pediatrics*, 129(1), e232-246. doi: 10.1542/peds.2011-2663
- Smith, L. K., Draper, E. S., Manktelow, B. N., Dorling, J. S., & Field, D. J. (2007). Socioeconomic inequalities in very preterm birth rates. *Archives of Disease in Childhood-Fetal and Neonatal Edition*, 92(1), F11-F14. doi: 10.1136/adc.2005.090308
- Voegtline, K. M., Costigan, K. A., Kivlighan, K. T., Laudenslager, M. L., Henderson, J. L., & DiPietro, J. A. (2013). Concurrent levels of maternal salivary cortisol are unrelated to self-reported psychological measures in low-risk pregnant women. *Archives of Women's Mental Health*, 16(2), 101-108. doi: 10.1007/s00737-012-0321-z
- Walker, L. O., & Avant, K. C. (2005). *Strategies for theory construction in nursing*.
- Wallace, M., Harville, E. (2013a). Allostatic Load and Birth Outcomes Among White and Black Women in New Orleans. *Maternal and Child Health Journal*, 17, 1025-1029.
- Wallace, M., Harville, E., Theall, K., Webber, L., Chen, W., & Berenson, G. (2013b). Neighborhood poverty, allostatic load, and birth outcomes in African American and white women: findings from the Bogalusa Heart Study. *Health & Place*, 24, 260-266. doi: <http://dx.doi.org/10.1016/j.healthplace.2013.10.002>
- Watson, L. F., & Taft, A. J. (2013). Intimate Partner Violence and the Association with Very Preterm Birth. *Birth-Issues in Perinatal Care*, 40(1), 17-23. doi: 10.1111/birt.12024
- Weiss, S. J. (2007). Neurobiological alterations associated with traumatic stress. *Perspectives in Psychiatric Care*, 43(3), 114-122. doi: 10.1111/j.1744-6163.2007.00120.x
- Xiong, X., Harville, E. W., Mattison, D. R., Elkind-Hirsch, K., Pridjian, G., & Buekens, P. (2008). Exposure to Hurricane Katrina, post-traumatic stress disorder and birth outcomes.

The American Journal of Medical Sciences, 336(2), 111-115. doi:
10.1097/MAJ.0b013e318180f21c

Yildiz, P. D., Ayers, S., Phillips, L. (2017). The prevalence of posttraumatic stress disorder in pregnancy and after birth: A systematic review and meta-analysis. *Journal of Affective Disorders*, 208, 634-645. doi: 10.1016/j.jad.2016.10.009

Yonkers, K. A., Smith, M. V., Forray, A., Epperson, C. N., Costello, D., Lin, H., & Belanger, K. (2014). Pregnant women with posttraumatic stress disorder and risk of preterm birth. *JAMA Psychiatry*, 71(8), 897-904. doi: 10.1001/jamapsychiatry.2014.558

Table 2-1: Brief descriptions of the 4 theoretical models

Models	Purposes	Main predictors	Moderators	Mediators	Outcomes
Beckie, 2012	To apply the AL theory to examine the dysregulated physiological stress responses and to predict diverse health outcomes.	Allostatic challenges including genetic, environmental, biographical, psychosocial, behavioral, and clinical challenges and childhood adversity	–	Primary mediators from neuroendocrine and immune/inflammatory system and secondary outcomes from metabolic and cardiovascular systems and inflammatory proteins	Mental and physical health, quality of life, and mortality
Seng, 2002	To describe the effects of lifetime violence and PTSD on childbearing women	Lifetime violence and PTSD	Non-modifiable factors and co-occurring risks, life event stress factors, and modifiable health care related factors	Behavioral and neuroendocrine alterations	Adverse maternal and fetal outcomes
Premji, 2014	To explicate the causal links of perinatal distress to preterm birth and infant health in low- and middle-income countries.	Episodic (e.g., traumatic life events, catastrophic events) and chronic stress (e.g., adverse socioeconomic, cultural, and environmental phenomena) and pregnancy-related mood disorders (e.g., depression, anxiety)	Prior context (e.g., childhood stress, history of depression, age)	AL	Pregnancy outcomes (i.e., still birth, spontaneous abortion, miscarriage, and preterm birth) and infant health

Olson et al., 2015	To understand the transgenerational links of cumulative life stressors to preterm birth.	Cumulative life stressors (i.e., early life adversity, social context, stress exposure)	Generational programming in connection with genetic and epigenetic signatures	Behavioral, psychological, and inflammatory stress responses; AL	Preterm birth
--------------------	--	---	---	--	---------------

Note. AL, allostatic load; PTSD, posttraumatic stress disorder.

Table 2-2: The new perinatal allostatic load model

Model	Purpose	Main predictors	Antecedents	Moderators	Mediators	Outcomes
New Proposed model	To understand pathway through which maternal PTSD results in adverse birth outcomes	PTSD	Chronic stress (i.e. CM, IPV, Low SES)	Depression, pregnancy-specific distress, and risk behaviors (i.e., tobacco, alcohol, and drug use)	Primary mediators from neuroendocrine system (e.g. cortisol, E, NE,), secondary outcomes from immune, metabolic, and cardiovascular systems, and anthropometric parameters, and tertiary outcomes (i.e. preeclampsia, GD, infection, and PROM)	Quaternary outcomes (preterm birth and low birth weight)

Note. PTSD, posttraumatic stress disorder; CM, child maltreatment; IPV, intimate partner violence; SES, socioeconomic status; E, epinephrine; NE, norepinephrine; GD, gestational diabetes; PROM, premature rupture of membrane.

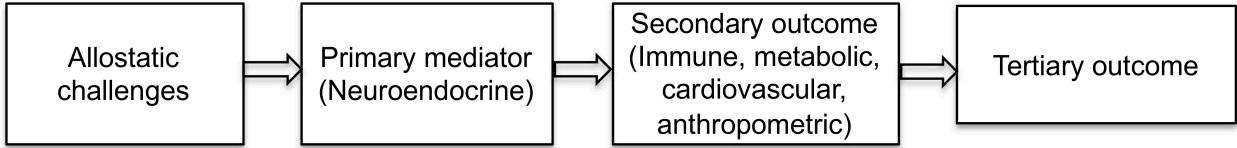


Figure 2-1: The process from stress to poor health outcomes based on the allostatic load theory
Note. The solid arrow indicates the causal relationship.

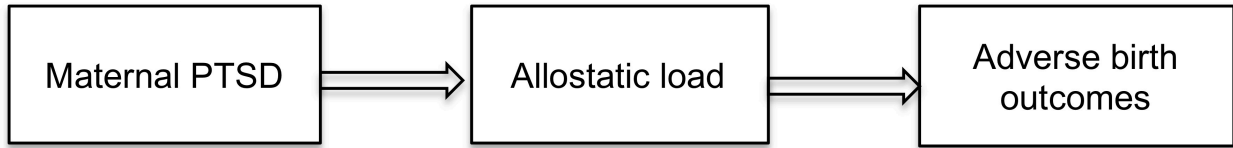


Figure 2-2: The mediating role of allostatic load in the link between maternal PTSD and adverse birth outcomes

Note. PTSD, posttraumatic stress disorder.

The solid arrow indicates the causal relationship.

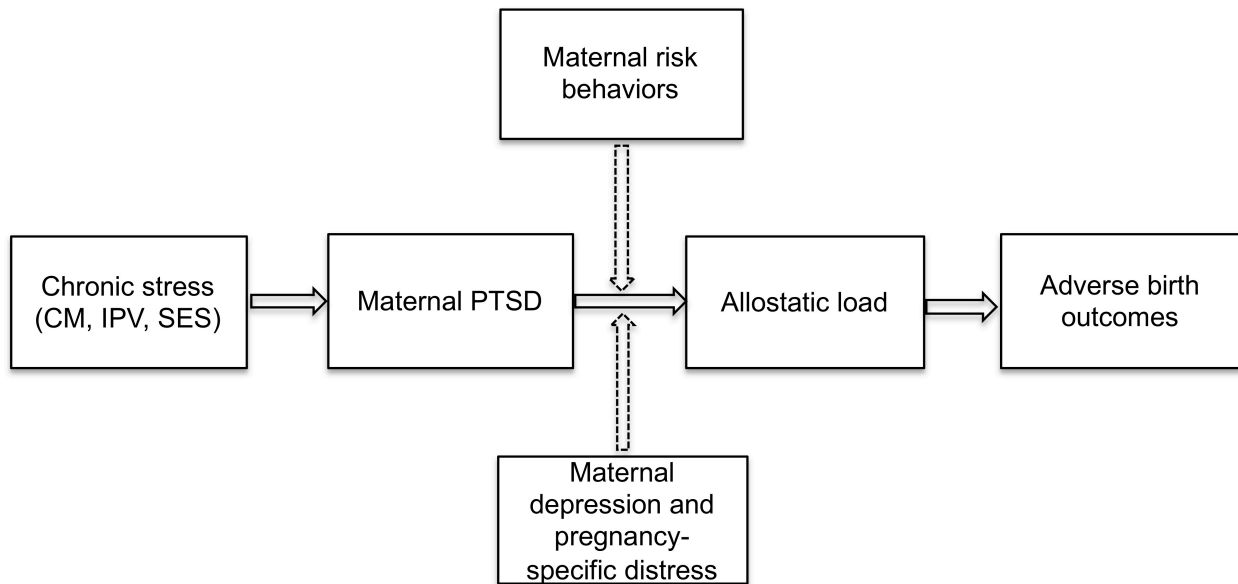
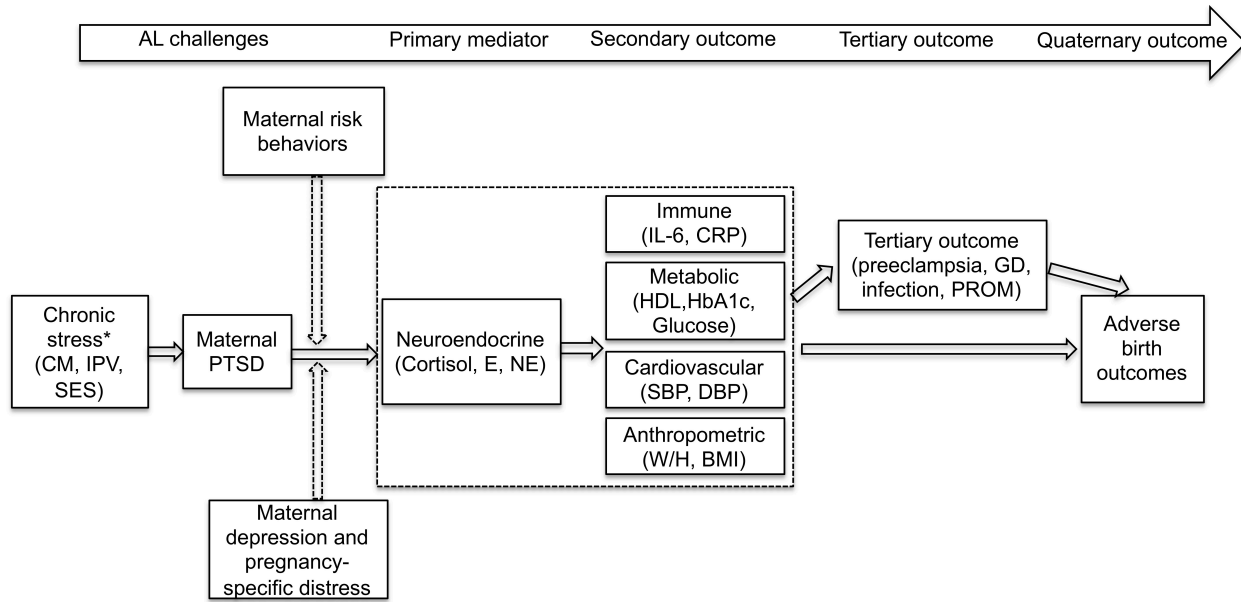


Figure 2-3: The influence of chronic stress, maternal risk behaviors, depression, and pregnancy-specific distress on the link between maternal PTSD and adverse birth outcomes

Note. CM, child maltreatment; IPV, intimate partner violence; SES, socioeconomic status; PTSD, posttraumatic stress disorder.

The solid arrow indicates the causal relationship; the dotted arrow indicates moderating role.



* Intergenerational patterns (i.e., genetics and epigenetics) may also play a role.

Figure 2-4: The theoretical model for understanding the link between maternal PTSD and adverse birth outcomes

Note. CM, child maltreatment; IPV, intimate partner violence; SES, socioeconomic status; PTSD, posttraumatic stress disorder; E, epinephrine; NE, norepinephrine; IL-6, interleukin-6; CRP, C-reactive protein; HDL, high density lipoprotein; HbA1c, glucosylated hemoglobin; SBP, systolic blood pressure; DBP, diastolic blood pressure; W/H, waist/hip; BMI, body mass index; GD, gestational diabetes; PROM, premature rupture of membrane.

The solid arrow indicates the causal relationship; the dotted arrow indicates moderating role.

Chapter 3: Exploring the optimal allostatic load scoring method in women of reproductive age

Abstract

Allostatic load refers to the accumulated physiological dysregulations across multiple systems in response to repeated, chronic stress, which may serve as a potential biological mechanism for understanding how chronic stress causes adverse health outcomes. But there is not yet a gold-standard scoring method of allostatic load that is valid across health outcomes. The study used data from the 2001-2006 National Health and Nutrition Examination Survey to determine the optimal scoring method by comparing 5 scoring methods including the count-based, Z-Score, logistic regression, factor analysis, and grade of membership method. We examined the predictive performance of each allostatic load summary measure with women of reproductive age in relation to 3 outcomes: general health status, diabetes, and hypertension. We found the allostatic load summary measure by the logistic regression method had the highest predictive validity with respect to the 3 outcomes. The logistic regression method performed significantly better than the count-based and grade of membership method for predicting diabetes as well as performed significantly better for predicting hypertension than all of the other methods. But the 5 scoring methods performed similarly for predicting poor health status. Each scoring approach has its own strengths and weaknesses. We recommended the logistic regression method when the outcome information is available, otherwise the frequently used, simpler count-based method

may be a good alternative.

Keywords: allostatic load, scoring, women of reproductive age

Introduction

Allostatic Load (AL) refers to the accumulated multi-system physiologic dysfunction resulting from repeated, chronic stress that could ultimately lead to disease (McEwen, 1998). When stress (e.g., socioeconomic disadvantage, child abuse and neglect) occurs, there is a cascade of effects that begins with primary stress mediators such as cortisol from the hypothalamic–pituitary–adrenal (HPA) axis, a primary effect, which in turn leads to secondary and tertiary outcomes (Beckie, 2012). The AL theory depicts how chronic stress leads to diseases. As a holistic measure of physiological dysfunction, AL may provide a multi-systemic approach to understand mechanisms involved in the impact of chronic stress on health.

AL is operationalized by combining physiological indicators from multiple systems (i.e., neuroendocrine, immune, metabolic, and cardiovascular) into one single index. The index is a more sophisticated, comprehensive physiological measure than a single system-specific indicator. It could reduce the probability of a type I error by combining multi-system indicators into one single index, rather than analyzing each individual indicator separately (McDade, 2008). However, there is no commonly accepted, gold-standard way to operationalize AL because of its multifaceted nature. Many scoring methods have been used to create an AL index (ALI) in previous studies, including the count-based, Z-Score, canonical correlation, recursive partitioning, and grade of membership (GOM) method. Controversies or challenges regarding AL scoring methods primarily arise from three issues: technique for calculating the index, weighting of respective indicators in the index, and norming on a population. Thus, the scoring

issue must be further considered before the concept of AL can be reliably and validly applied to research and clinical practice.

Current scoring approaches

The most frequently used scoring method is the count-based method. The ALI by this method is the sum of the number of indicators for which an individual falls into the risk quartile of the sample distribution (Seeman, Singer, Rowe, Horwitz, & McEwen, 1997). The risk quartile has been commonly used as a cut-off point for indicators. It has been also found that the differences in predictive performance were not substantial when the quartile cut-off point is compared with the risk decile (Hampson, Goldberg, Vogt, Hillier, & Dubanoski, 2009). It is simple to calculate the overall index using the count-based method, but dichotomizing each individual indicator would lose information regarding the potential variability in their contributions to overall risk and might decrease the statistical power in analyses (Seeman et al., 2008). This method also has the limitation of making the ALI sample-specific by dichotomizing indicators based on the risk quartile of the sample distribution. For all AL indicators, no current population norms in terms of age, race, sex, etc. have ever been derived. Thus, the sample-specific summary measure may not be meaningfully compared across samples. Furthermore, all physiological indicators count equally in the summary score. The relative importance of various physiological components to the overall score for predicting health outcomes is not considered. Some indicators may be more critical than others with regard to certain outcomes.

Another relatively simple scoring approach is the Z-Score method. In this approach, all indicators are individually standardized to a mean of zero and a standard deviation of one. The ALI is the sum of the standardized distance of each indicator from its respective mean. The formulation is based on a continuous, rather than a categorical, function of the biological

measures (Vie, Hufthammer, Holmen, Meland, & Breidablik, 2014). Compared to the count-based method, the Z-Scored ALI could account for some variances and increase the statistical power (Hampson et al., 2009). But it is still sample-specific and fails to account for the weighting of each indicator in the summary measure.

Some AL studies have applied other scoring methods that are more complex than a simple count or a Z score, such as canonical correlation, recursive partitioning, and GOM. These alternative scoring approaches provide more complex scoring algorithms and incorporate more information of each individual indicator than the simple counting of high-risk cut-off points. They also allow for unequal weighting of various biological measures (Beckie, 2012). Canonical correlation has been used to determine the best linear combination of AL indicators that is maximally correlated with the best linear combination of health outcomes (Karlmanjla, Singer, McEwen, Rowe, & Seeman, 2002). An AL summary score can be constructed using the sets of AL indicators and their canonical weights in the best linear correlation. This approach permits unequal weights for each AL indicator, but it requires continuous variables and relies on the subsequent outcome information. Since the canonical weights are derived from and applied to the same sample, it makes the ALI too specific to the data used to derive it. This may magnify the predictive ability of the index, deplete its predictive power in other contexts, and cause the endogeneity bias, that is to say it may not be generalized to other contexts (Seplaki, Goldman, Gleib, & Weinstein, 2005a). Recursive partitioning is a technique that has been used to classify individuals into outcome risk categories. It can identify multiple combinations of physiological indicators and their value ranges to best differentiate among outcomes across individuals (Juster, McEwen, & Lupien, 2010). It was also used to define AL categories (e.g., high, intermediate, low). Similar to the canonical correlation, this approach has the limitation of incorporating

information on subsequent health outcomes (Seplaki et al., 2005a). The GOM method was used to create pre-defined pure profiles (N), which are the collections of response probabilities corresponding to each level of discrete indicators. Accordingly, N GOM (summing to one) scores were assigned to each individual, measuring the similarity of the set of a person's indicator values to each respective profile. The GOM score-based ALI is the sum of N-1 of the GOM scores (excluding the score for the reference/low risk profile), measuring dissimilarity to the low risk profile (Seplaki, Goldman, Weinstein, & Lin, 2006). The method does not incorporate information on subsequent health outcomes, but still categorizes each indicator into low, moderate, or high levels based on the sample distribution.

Three prior studies used factor analysis to construct and evaluate structural models of AL reflecting the cumulative physiological burden across multiple systems (Booth, Starr, & Deary, 2013; Kubzansky, Kawachi, & Sparrow, 1999; Seeman et al., 2010). Parameter estimates obtained from factor analysis can be considered as the specific contributions of respective indicators to the summary score. Studies on creating other clinical index measures have used some other statistical techniques such as the multivariable logistic regression (Hughes et al., 2012; Lee, Lindquist, Segal, & Covinsky, 2006). The multivariable logistic regressions were fitted with all potential components as predictors and outcomes as dependent variables. Coefficients obtained from the regression models can be considered as weights for each component. Scores were allocated to each component based on those weights and summed up to a total index. But to our knowledge, no previous studies have used the factor analysis or logistic regression method to assign weights to each AL indicator.

Despite continued work using various scoring methods to construct the ALI, there is not yet a gold-standard measure of AL that is valid across health outcomes. No studies focused on

comparing different scoring methods and determining which method is optimal to capture the multiple and inter-connected feature of AL and summarize such information into one cumulative index. Such comparative data remains sparse and the question of how best to score AL remains to be addressed. The lack of comparative analyses of the reliability of different AL measures represents a gap in the current research on chronic stress and health outcomes. Thus, more comparative work is clearly needed to compare the explanatory powers and predictive abilities of distinct scoring approaches and to determine the optimal scoring approach before examining AL as a mediating pathway for the impact of chronic stress on health outcomes.

Study aims

This study aimed to determine the optimal AL scoring method by comparing several scoring methods within a single population dataset. Because age and gender would influence the AL summary score, the study focused on a more homogeneous female population – women of reproductive age from the 2001-2006 National Health and Nutrition Examination Survey (NHANES) database. We constructed the ALI using 5 scoring methods including the count-based, Z-Score, logistic regression, factor analysis, and GOM method. We then examined the predictive performance of each ALI with women of reproductive age in relation to 3 outcomes: self-reported general health status, diabetes, and hypertension. None of the scoring methods we examined incorporate these outcomes in the calculation of the summary measure except the logistic regression method. The logistic regression, factor analysis, and GOM method considered the weighting issue. Because of the limitations of the NHANES database, we excluded some scoring methods. The canonical correlation approach was excluded because it requires continuous variables including outcome variables, while the outcome variables available in the NHANES database are categorical. The recursive partitioning technique was not used in the

study because only AL categories (e.g., high, intermediate, low) can be defined and no total score can be constructed through this approach.

Methods

Study design

This is a secondary analysis of data from the NHANES. NHANES is a cross-sectional study with a complex, multistage probability sampling design used to select a sample representative of the civilian non-institutionalized resident population of the United States, which has been conducted in 2-year cycles since 1999 (Curtin et al., 2012). In this study, we used data from the 2001 to 2006 cycles of NHANES to test the study aims. Data from 2007-2010 were used to replicate the main analyses and compare the results with the 2001-2006 data to evaluate the stability of the results. The data collected between 1999-2000 were not used because general health status was not queried during that 2-year cycle. Data collected after 2010 was not used because no C-reactive protein (CRP) has been measure since 2011. The NHANES 2001–2010 were approved by the National Center for Health Statistics Research Ethics Review Board under protocols #98-12 and #2005-06 and Continuation of Protocol #2005-06. This secondary analysis of data was exempt from IRB review because it was done via the de-identified dataset.

Participants

Female participants with reproductive ages of 15-49 were included in the study. Women who were pregnant at the exam measured by the urine pregnancy test were excluded. A total of 5525 women were eligible for the study in the 2001-2006 NHANES data. But 1206 women (21.8%) had missing data on the 3 outcome variables (general health status, diabetes, and hypertension). Thus, 4319 women were finally included for analysis to address the study aims. In

the 2007-2010 NHANES data, a total of 3018 women were included to replicate the main analyses.

Variables and data sources

AL. The selected 10 indicators in this study were CRP, systolic blood pressure (SBP), diastolic blood pressure (DBP), pulse, body mass index (BMI), total cholesterol (TC), high-density lipoprotein (HDL), triglycerides, glycohemoglobin, and glucose. These indicators were frequently used in previous studies (Juster et al., 2010). Other indicators in the NHANES database were not included in the study because there is a large amount of missing data or some of those indicators were collected only in subsamples. Standard examination and laboratory procedures were described in the NHANES Examination and Laboratory Protocols (CDC & NCHS).

Outcomes. General health status was measured by using 1 item asking whether participants' general health is excellent, very good, good, fair, or poor from the current health status questionnaire. In this study, it was dichotomized into two levels: "poor" and "fair, good, very good, or excellent". We used 1 item— "Other than during pregnancy, have you ever been told by a doctor or health professional that you have diabetes or sugar diabetes?" from the diabetes questionnaire to determine diabetes being present or not. Participants who reported "Borderline" were considered as no diabetes. The question—"Have you ever been told by a doctor or other health professional that you had hypertension, also called high blood pressure?" from the blood pressure & cholesterol questionnaire was used to determine hypertension present or not.

Sociodemographic characteristics. Age, race, poverty income ratio (PIR), education, and marital status from the demographic dataset were included in this study. Age was measured

as age in months at the time of examination. PIR is an index for the ratio of family income to poverty threshold, ranging between 0-5.00. We dichotomized race into 2 categories: non-Hispanic black and other races (e.g., Mexican American, other Hispanic, non-Hispanic white, and other race including multi-racial). Education level was categorized into: Less than high school, high school diploma including GED, and more than high school. Marital status was recoded as married/living with partner and widowed/divorced/separated/never married.

AL scoring methods

The count-based method. A dichotomous high-risk score was computed for each indicator by assigning a score of 1 to participants whose scores were within the top risk quartile of the sample distribution (75th percentile for all indicators except HDL for which 25th percentile corresponds to high risk) and a score of 0 otherwise. An ALI was then constructed as the sum of the 10 dichotomous (0/1) indicator risk scores, yielding a possible score range of 0-10.

The Z-Score method. All 10 indicators were individually standardized to a mean of zero and a standard deviation of one. The HDL Z-Score was reversed so that high values reflect greater dysregulation. An ALI was then calculated by summing the Z-Scores of all indicators.

The logistic regression method. Multivariable logistic regressions were conducted with all 10 AL indicators as independent variables and the 3 outcome variables (i.e., general health status, diabetes, and hypertension) as the dependent variable respectively. The standardized coefficients obtained from the models were used as the weights for each individual indicator. Indicators were first individually standardized to a mean of zero and a standard deviation of one. The Z-Scores were then multiplied by the coefficients for each individual indicator derived from

the regression models. Using this method with the 3 outcome variables as the dependent variable respectively, 3 ALI were computed by summing the multiplied values for each indicator.

The factor analysis. We conducted the factor analysis using robust maximum likelihood estimation with the number of factors set as 1. Indicators were first individually standardized to a mean of zero and a standard deviation of one. The Z-Scores were multiplied by the factor loading for each individual indicator derived from the factor analysis. We then created the summation scores for ALI.

The GOM. Each indicator was divided into low and high risk for poor health based on the 75th percentile of the sample distribution except HDL for which 25th percentile was the risk quartile. The number of pure-type profiles was set in advance. Each pure-type profile is a collection of response probabilities corresponding to each level of the 10 discrete indicators. Our analyses showed that compared to 3 and 4 pure types, 5 pure-type profiles provide reasonable interpretability and summaries of the physiological functions. Detailed definitions for the 5 pure types can be seen in Figure 3-1A and 3-1B. Accordingly, a set of 5 GOM scores for each individual that quantify the individual's similarity to each pure-type profile was created, ranging from 0-1 and summing to unity. Excluding the score measuring similarity to the low-risk, or reference, pure-type profile (the 5th profile), the other 4 GOM scores were summed to create a single GOM-based AL summary measure, reflecting dissimilarity to the low-risk profile. Detailed explanations for the GOM method can be found in previous studies (Seplaki et al., 2005a; Seplaki, Goldman, Weinstein, & Lin, 2004; Seplaki et al., 2006).

Among the 10 indicators, glucose and glycohemoglobin are direct clinical indicators for the diagnosis of diabetes and SBP and DBP are directly related with the diagnosis of hypertension. An issue that arises is whether the associations between ALI and diabetes or

hypertension reflect only or largely the impact of the 4 indicators or whether the other indicators have significant and independent relationships with these two outcomes. Thus, using the 5 scoring approaches, we also constructed tailored ALI without glucose and glycohemoglobin predicting diabetes and without SBP and DBP predicting hypertension. Results based on tailored ALI (8 indicators) were compared with the results of ALI (10 indicators) in a sensitivity analysis.

Statistical analysis

Means, standard deviations, 25th/75th percentiles, frequencies, and percentages were used to describe sociodemographic characteristics, the 3 outcome variables, and the 10 AL indicators. The multiple imputation (MI) method (Rubin, 2004) was used to impute all missing data. We used chained equations and predictive mean matching with non-missing sociodemographic variables and indicators as predictor variables. The imputations of the missing values are predicted values from these regression models, with the appropriate random error included. Since there is 17.6% of data missing, 10 imputed datasets were created. In each of the imputed datasets, we conducted all main analyses including constructing the ALI with different scoring approaches and validating the index. The overall estimate is the average of the estimates from each of the imputed datasets.

The distributional qualities, including range, mean, standard deviation, median, skew, and kurtosis, were used to describe AL summary measures by each of the 5 scoring methods. The odds ratio (OR) by each method was computed through fitting binomial logistic regression models to estimate the strengths of the associations of each AL summary measure with general health status, diabetes, and hypertension respectively. The 3 outcomes were included as the dependent variable respectively and each summary measure of AL was included as the independent variable. The covariates included age, race, and PIR. All ALI scores by the 5

methods were standardized to a mean of zero and a standard deviation of one before fitting the regression models, so that the strengths of the (adjusted) associations between AL summary measures and outcomes can be compared across different scoring approaches. Additionally, the areas under the receiver operating characteristic (ROC) curve (AUC) were calculated to estimate the predictive validity of each AL summary measure for predicting the 3 outcomes. An AUC with successively higher values above 0.5 indicates increasing levels of predictive value.

To investigate the performance of different AL measures in an external sample, the process was subsequently repeated, conducting the same analyses in the NHANES 2007-2010 dataset. In order to make a recommendation of the optimal scoring method for clinical use purposes, we also evaluated each scoring method by qualitative comparisons in terms of strengths and weaknesses. Using the optimal scoring method, we calculated the cut-off points, sensitivities, and specificities. All statistical analyses were performed using R Software Version 3.4.2 (R Core Team, 2017).

Results

The sample characteristics

The mean age of the sample was approximately 30 years and about 26% of women were non-Hispanic Black. Around 58% reported completing high school or less than high school education and 56.5% were married or living with partner. Only 1.7% reported poor health status, 3.1% had diabetes, and 12.6% had hypertension (Table 3-1). Table 3-2 showed the descriptive statistics of each AL indicator.

The descriptive statistics of ALI

The ALI constructed by the count-based and GOM method did not consist of negative values, while the ALI ranged from a negative value to a positive value for the other 3 methods.

The skew and kurtosis of the ALI using the count-based method, the logistic regression with general health and diabetes as the outcome, and the GOM method were close to 0, indicating these indices are more normally distributed (Table 3-3). The skew and kurtosis of the tailored ALI using the count-based measure, tailored ALI without glucose and glycohemoglobin using the logistic regression, and tailored ALI without SBP and DBP using the GOM were less than 1, suggesting the distributions of those indices were more normal (Table 3-4). The distributions of each ALI and tailored ALI were shown in Figure 3-2 and 3-3. All distributions were unimodal except for the tailored ALI without glucose and glycohemoglobin using the GOM method. Interestingly, the tailored ALI without glucose and glycohemoglobin using the GOM method presented a bimodal distribution with two peaks close to 0 and 1 respectively, which visually showed the cut-off point of the ALI for poor health risk.

The predictive validities of ALI

The logistic regression method was most strongly associated with the 3 outcome measures, whether adjusted or not adjusted (Table 3-5). This remained the case when 2 indicators diagnostic for diabetes or hypertension were removed from the index (Table 3-6). Using the factor analysis method, the associations of ALI with general health and hypertension were smallest and significantly smaller than the logistic regression method. But there were no significant differences in terms of the strengths of the associations among the count-based, Z-Score, logistic regression, and GOM method. The count-based measure was nearly as strongly related to the outcome measures as the logistic regression, adjusted or unadjusted, tailored or not. As expected, all ALI with 10 indicators were more strongly associated with diabetes and hypertension compared to the tailored ALI without glucose and glycohemoglobin and tailored ALI without SBP and DBP.

The 5 scoring methods had similar predictive performances with regard to general health. But the logistic regression method had better predictive performance for predicting diabetes compared to the count-based and GOM method and had the best performance for predicting hypertension than the other 4 methods. The ALI by any method predicted diabetes and hypertension better than it predicted the subjective appraisal of overall health status (Table 3-7 and Figure 3-4 and 3-6). The tailored ALI (excluding glucose and glycohemoglobin or SBP and DBP) by any method had similar predictive validities in terms of diabetes and hypertension except that the logistic regression method predicted hypertension better than the GOM method. As expected, the tailored ALI by any method had worse predictive powers compared to the ALI with all 10 indicators included (Table 3-8 and Figure 3-5 and 3-7).

Parallel analyses

All analyses were conducted again using the NHANES 2007-2010 data, yielding approximately the same results. Similarly, the logistic regression method had the strongest associations with the outcome measures, whether adjusted or not adjusted, tailored or not. The count-based method was nearly as strongly associated with the outcome measures as the logistic regression, adjusted or unadjusted, tailored or not. The 5 scoring methods had similar predictive validities with regard to the 3 outcome measures. Similarly, the logistic regression method still had the best predictive performances, whether tailored or not.

Discussion

This study constructed an ALI using 5 scoring approaches and assessed the predictive performances across different scoring approaches in women of reproductive age. We found the AL summary measure by the logistic regression method had the strongest predictive validity with respect to general health status, diabetes, and hypertension. The logistic regression method

performed significantly better than the count-based and GOM method for predicting diabetes as well as performed significantly better for predicting hypertension than all of the other methods. But the 5 scoring methods performed similarly for predicting poor health status. Excluding the diagnostic indicators for diabetes and hypertension, the independent contributions of the other 8 indicators to the risk of diabetes and hypertension were demonstrated. Differences in the predictive performances in terms of diabetes and hypertension became smaller among the 5 scoring methods, but the logistic regression method still performed the best. The findings were duplicated using the 2007-2010 NHANES data, underscoring the robustness of the finding.

The predictive performances across different scoring methods in this study are similar, which is partially consistent with a study using data from a population-based sample of older Taiwanese to compare several count-based formulations as well as the Z-Score and GOM method. All AL summary measures had similar predictive performances for predicting self-assessed health, impairments in activities of daily living (ADL) and mobility, cognitive performance, and depressive symptoms. The study recommended the count-based and Z-Score method since the two methods are simple to compute and the GOM method is more complicated (Seplaki, Goldman, Gleib, & Weinstein, 2005b). Another study with a community sample of 470 participants from the Hawaii Personality and Health cohort also reported similar performances of the count-based and Z-Score method for predicting self-rated health (Hampson et al., 2009). The differences among the 5 summary measures were not pronounced in this study, suggesting that the advantages of one method over another are relatively subtle.

The differences in the predictive performances between the logistic regression method and the other scoring methods for predicting diabetes and hypertension were larger than for predicting poor health status. In addition, the differences became smaller after excluding the 4

diagnostic indicators (glucose, glycohemoglobin, SBP, and DBP) for diabetes and hypertension. Given that the logistic regression method accounts for the non-uniform contributions of distinct biological measures to health risk, the possible explanation for this finding is that large weightings were assigned to the 4 diagnostic indicators by the logistic regression method. The finding suggests that the logistic regression method predicts better when some AL components have much stronger associations with specific health outcomes than the other AL components.

Each scoring approach has its own strengths and weaknesses (Table 3-9). The ALI by the logistic regression method had the best predictive performance compared to the other methods. But this method assigns scoring weights to each indicator based on information on subsequent outcomes. It is challenging to compare AL summary scores across different outcomes. And the logistic regression method may not be the optimal scoring method when the outcome information is unknown. For example, in the preliminary stage of a research project, only data on physiological indicators is available while data on the targeted outcome has not been collected. Also, the outcome is not needed for some studies that only focus on exploring some stressors in relation to AL levels.

Under the above conditions, the count-based method may be a good alternative. The predictive performance of the ALI by the count-based method for predicting general health status is similar to the other approaches and even for predicting diabetes and hypertension is similar to the other approaches except the logistic regression method. Additionally, after excluding the diagnostic indicators for diabetes and hypertension, the count-based method performed as well as the logistic regression method for predicting diabetes and hypertension. Compared to the other methods, the count-based method has its own strengths. It is the most frequently used method in prior AL studies. The AL summary score by this method is the number of indicators of risk for

poor health, which is a real value and easy to interpret. It is simple to calculate, easy to understand, and feasible to be applied in clinical practice. Therefore, if the outcome information is available, needed, and consistent across different contexts, we recommended the logistic regression method; otherwise the count-based method may be a good alternative.

Using the count-based method, we calculated cut-off points, sensitivities, and specificities of the ALI score (Table 3-10). Although the count-based method had some strengths, it has the limitation of making the ALI sample-specific. A better way to address the limitation is to use the clinical risk cut-off points based on national standards instead of risk quartiles of the sample distribution to count the total number of indicators of risk for poor health. But further work on establishing population norms in terms of age, sex, race, etc. is needed. Especially, no current population norms for pregnancy have ever been derived, which make it challenging to apply the AL theory to perinatal outcomes research.

This study had some limitations. First, we focused on women of reproductive age. Because the dysregulated levels of each AL indicator are different in terms of age and gender, scoring AL in a more homogeneous female population may contribute to the reliability of our findings. But our findings may not be generalized to the male or elder population. Future research needs to replicate our analyses in different age- and gender-specific populations. Age- and gender-specific population norms for the ALI score by the optimal scoring method will be also needed. Second, data on indicators from the primary mediating neuroendocrine system are lacking in the NHANES database. The ALI was constructed without indicators from the neuroendocrine system, relying solely in the indicators of secondary dysregulations for the scoring, which may decrease the predictive validity and explanatory power of the total score on health outcomes. Third, because of the cross-sectional study design of the NHANES, data on the

outcome variables and AL indicators were collected at the same time. This may also affect the predictive performances of ALI for predicting general health status, diabetes, and hypertension. A prospective study using a full complement of physiological indicators to operationalize the AL and using different scoring approaches is needed to validate the recommendations made based on this secondary analysis. The study has strengths as well. The NHANES database includes numerous physiological measures from multiple body systems, which provides adequate AL indicators for our study to create the ALI. Additionally, data are collected with standardized procedures and protocols to assure that the data for this analysis is of high quality in terms of reliability and validity.

In summary, the logistic regression method is the optimal scoring method for use, especially in research. The count-based method may be a good alternative for research and clinical practice when the outcome information is not available/needed. Our study advanced studies of AL by focusing on scoring methods with a nationally representative dataset and making recommendations for the optimal method to score AL. It has potential implications for research and clinical practice. It provides empirical evidence for researchers to use the recommended scoring approach to score AL in their research. Our findings may also be useful for clinicians. The ALI can serve as a sign for risk of subclinical syndromes. Most of AL indicators such as BMI, blood pressure, and pulse are routine clinical assessments and thus are feasible to be measured. The logistic regression method can be used through computer software and the count-based method as an alternative measure can be easily calculated by hand. Therefore, the AL summary measure is easy and feasible for use as an “early warning” indicator for health risk across a variety of care settings.

References

- Beckie, T. M. (2012). A systematic review of allostatic load, health, and health disparities. *Biological Research for Nursing, 14*(4), 311-346.
- Booth, T., Starr, J. M., & Deary, I. (2013). Modeling multisystem biological risk in later life: allostatic load in the Lothian birth cohort study 1936. *American Journal of Human Biology, 25*(4), 538-543. doi: <http://dx.doi.org/10.1002/ajhb.22406>
- Curtin, L. R., Mohadjer, L. K., Dohrmann, S. M., Montaquila, J. M., Kruszan-Moran, D., Mirel, L. B., . . . Johnson, C. L. (2012). The National Health and Nutrition Examination Survey: Sample Design, 1999-2006. *Vital and Health Statistics, 2*(155), 1-39.
- Hampson, S. E., Goldberg, L. R., Vogt, T. M., Hillier, T. A., & Dubanoski, J. P. (2009). Using Physiological Dysregulation to Assess Global Health Status Associations with Self-rated Health and Health Behaviors. *Journal of Health Psychology, 14*(2), 232-241. doi: 10.1177/1359105308100207
- Hughes, D. A., Malmenas, M., Deegan, P. B., Elliott, P. M., Ginsberg, L., Hajioff, D., . . . Investigators, F. O. S. (2012). Fabry International Prognostic Index: a predictive severity score for Anderson-Fabry disease. *Journal of Medical Genetics, 49*(3), 212-220. doi: 10.1136/jmedgenet-2011-100407
- Juster, R. P., McEwen, B. S., & Lupien, S. J. (2010). Allostatic load biomarkers of chronic stress and impact on health and cognition. *Neuroscience & Biobehavioral Reviews, 35*(1), 2-16. doi: <http://dx.doi.org/10.1016/j.neubiorev.2009.10.002>
- Karlamangla, A. S., Singer, B. H., McEwen, B. S., Rowe, J. W., & Seeman, T. E. (2002). Allostatic load as a predictor of functional decline. MacArthur studies of successful aging. *Journal of Clinical Epidemiology, 55*(7), 696-710.
- Kubzansky, L. D., Kawachi, I., & Sparrow, D. (1999). Socioeconomic status, hostility, and risk factor clustering in the Normative Aging Study: any help from the concept of allostatic load?. *Annals of Behavioral Medicine, 21*(4), 330-338.
- Lee, S. J., Lindquist, K., Segal, M. R., & Covinsky, K. E. (2006). Development and validation of a prognostic index for 4-year mortality in older adults. *JAMA, 295*(7), 801-808.
- McDade, T. W. (2008). Challenges and opportunities for integrative health research in the context of culture: A commentary on Gersten. *Social Science & Medicine, 66*(3), 520-524. doi: 10.1016/j.socscimed.2007.09.005
- McEwen, B. S. (1998). Protective and damaging effects of stress mediators. *The New England Journal of Medicine, 338*(3), 171-179. doi: 10.1056/NEJM199801153380307
- Rubin, D. B. (2004). *Multiple imputation for nonresponse in surveys* (Vol. 81): John Wiley & Sons.
- Seeman, T., Gruenewald, T., Karlamangla, A., Sidney, S., Liu, K., McEwen, B., & Schwartz, J. (2010). Modeling multisystem biological risk in young adults: The Coronary Artery Risk Development in Young Adults Study. *American Journal of Human Biology, 22*(4), 463-472. doi: <http://dx.doi.org/10.1002/ajhb.21018>

- Seeman, T., Merkin, S. S., Crimmins, E., Koretz, B., Charette, S., & Karlamangla, A. (2008). Education, income and ethnic differences in cumulative biological risk profiles in a national sample of US adults: NHANES III (1988-1994). *Social Science & Medicine*, *66*(1), 72-87. doi: 10.1016/j.socscimed.2007.08.027
- Seeman, T. E., Singer, B. H., Rowe, J. W., Horwitz, R. I., & McEwen, B. S. (1997). Price of adaptation--allostatic load and its health consequences. MacArthur studies of successful aging. *Archives of Internal Medicine*, *157*(19), 2259-2268.
- Seplaki, C. L., Goldman, N., Gleib, D., & Weinstein, M. (2005a). A comparative analysis of measurement approaches for physiological dysregulation in an older population. *Experimental Gerontology*, *40*(5), 438-449.
- Seplaki, C. L., Goldman, N., Gleib, D., & Weinstein, M. (2005b). A comparative analysis of measurement approaches for physiological dysregulation in an older population. *Experimental Gerontology*, *40*(5), 438-449. doi: 10.1016/j.exger.2005.03.002
- Seplaki, C. L., Goldman, N., Weinstein, M., & Lin, Y. H. (2004). How are biomarkers related to physical and mental well-being? *The Journals of Gerontology: Series A*, *59*(3), 201-217.
- Seplaki, C. L., Goldman, N., Weinstein, M., & Lin, Y. H. (2006). Measurement of cumulative physiological dysregulation in an older population. *Demography*, *43*(1), 165-183.
- R Core Team (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. .
- Vie, T. L., Hufthammer, K. O., Holmen, T. L., Meland, E., & Breidablik, H. J. (2014). Is self-rated health a stable and predictive factor for allostatic load in early adulthood? Findings from the Nord Trondelag Health Study (HUNT). *Social Science & Medicine*, *117*, 1-9. doi: 10.1016/j.socscimed.2014.07.019

Table 3-1: The descriptive statistics of sample sociodemographics and health outcomes
(*N*=4319)

	<i>N</i>	<i>M</i> ± <i>SD</i> / <i>%</i>
Age	4319	29.58±10.76
Poverty income ratio	4112	2.40±1.64
Race	4319	
Mexican American	1038	24.03
Other Hispanic	179	4.14
Non-Hispanic White	1786	41.35
Non-Hispanic Black	1130	26.16
Other Race - Including Multi-Racial	186	4.31
Education level	4317	
Less than high school	1609	37.27
High school diploma including GED	897	20.78
More than high school	1811	41.95
Marital status	4318	
Married/living with partner	2439	56.48
Widowed/divorced/separated/never married	1879	43.52
General health	4319	
Excellent	484	11.21
Very good	1427	33.04
Good	1667	38.60
Fair	666	15.42
Poor	75	1.7
Diabetes	4319	
Yes	133	3.08
No	4186	96.92
Hypertension	4319	
Yes	546	12.64
No	3773	87.36

Table 3-2: The descriptive statistics of the 10 allostatic load indicators ($N=4319$)

	<i>N</i>	<i>M</i> ± <i>SD</i> / <i>%</i>	Percent 25th/75th ^a
Pulse, beat per min	4225	76.13±11.57	84.0
SBP, mmHg	4183	111.87±13.19	118.0
DBP, mmHg	4023	67.78±10.87	74.5
BMI	4263	27.52±7.48	31.73
TC, mg/dL	4060	183.14±38.19	205.0
HDL, mg/dL	4060	55.67±14.66	45.0
CRP, mg/dL	4089	0.42±0.78	0.48
Glycohemoglobin, %	4116	5.27±0.69	5.4
Glucose, mg/dL	4056	88.39±20.48	91.0
Triglycerides, mg/dL	4056	104.81±85.08	126.0

Note. SBP, systolic blood pressure; DBP, diastolic blood pressure; BMI, body mass index; TC, total cholesterol; HDL, high-density lipoprotein; CRP, C-reactive protein

^a The percent 75th was calculated for all AL indicators except HDL for which the percent 25th was calculated.

Table 3-3: The descriptive statistics of allostatic load indices using the 5 scoring methods

	$M \pm SD$	Median	Min-Max	Skew	Kurtosis
Count-based method	2.35±2.03	2	0-10	0.91	0.31
Z-Score method	0±4.89	-0.92	-10.53-38.11	1.62	5.52
Logistic regression					
General health as the outcome	0±0.81	-0.11	-2.48-4.36	0.76	1.08
Diabetes as the outcome	0±1.41	-0.23	-5.36-17.00	4.33	32.67
Hypertension as the outcome	0±1.06	-0.16	-4.28-5.56	0.86	1.44
Factor analysis	0±0.62	-0.12	-1.03-7.46	4.47	32.88
Grade of membership	0.30±0.29	0.22	0.02-0.94	0.71	-0.74

Table 3-4: The descriptive statistics of tailored allostatic load indices (8 indicators) using the 5 scoring methods

	<i>M±SD</i>	Median	Min-Max	Skew	Kurtosis
Count-based method					
Tailored ALI without glucose and glycohemoglobin	1.93±1.70	2	0-8	0.83	0.06
Tailored ALI without SBP and DBP	1.87±1.70	1	0-8	0.93	0.32
Z-Score method					
Tailored ALI without glucose and glycohemoglobin	0±3.98	-0.61	-9.38-32.49	0.99	2.39
Tailored ALI without SBP and DBP	0±4.14	-0.75	-8.45-35.38	2.04	8.68
Logistic regression					
Tailored ALI without glucose and glycohemoglobin ^a	0±0.88	-0.15	-2.13-4.56	0.84	0.94
Tailored ALI without SBP and DBP ^b	0±0.71	-0.15	-1.29-4.36	1.23	2.18
Factor analysis					
Tailored ALI without glucose and glycohemoglobin	0±0.55	-0.10	-1.31-3.76	1.07	2.04
Tailored ALI without SBP and DBP	0±0.67	-0.12	-1.03-8.39	5.11	40.53
Grade of membership					
Tailored ALI without glucose and glycohemoglobin	0.44±0.45	0.30	0.007-0.99	0.27	-1.75
Tailored ALI without SBP and DBP	0.26±0.32	0.017	0.017-0.94	0.89	-0.63

Note. ALI, allostatic load index; SBP, systolic blood pressure; DBP, diastolic blood pressure.

^a The tailored ALI without glucose and glycohemoglobin was constructed using the logistic regression with diabetes as the outcome variable;

^b The tailored ALI without SBP and DBP was constructed using the logistic regression with hypertension as the outcome variable.

Table 3-5: The binary logistic regressions of allostatic load indices by the 5 scoring methods on general health, diabetes, and hypertension

	General Health		Diabetes		Hypertension	
	OR (95% CI)	Adjusted OR (95% CI)	OR (95% CI)	Adjusted OR (95% CI)	OR (95% CI)	Adjusted OR (95% CI)
Count-based method	2.11 (1.74-2.55)	1.68 (1.35-2.10)	3.15 (2.69-3.70)	2.67 (2.24-3.19)	2.32 (2.12-2.54)	1.90 (1.72-2.10)
Z-Score method	1.84 (1.59-2.14)	1.53 (1.29-1.82)	3.42 (2.91-4.03)	3.12 (2.62-3.71)	2.19 (1.99-2.41)	1.81 (1.64-2.00)
Logistic regression	2.26 (1.87-2.73)	1.86 (1.50-2.30)	4.10 (3.41-4.92)	3.68 (3.07-4.43)	2.88 (2.60-3.19)	2.34 (2.09-2.62)
Factor analysis	1.43 (1.29-1.59)	1.25 (1.11-1.42)	3.61 (3.05-4.29)	3.27 (2.75-3.90)	1.84 (1.67-2.03)	1.48 (1.35-1.62)
Grade of membership	2.06 (1.66-2.57)	1.62 (1.28-2.05)	3.27 (2.69-3.98)	2.71 (2.21-3.33)	2.04 (1.86-2.23)	1.71 (1.55-1.89)

Note. Age, race, and poverty income ratio were adjusted for.

Table 3-6: The binary logistic regressions of tailored allostatic load indices by the 5 scoring methods on diabetes and hypertension

	Diabetes		Hypertension	
	OR	Adjusted OR	OR	Adjusted OR
Count-based method	2.34 (2.00-2.73)	1.90 (1.61-2.25)	1.89 (1.73-2.06)	1.57 (1.43-1.73)
Z-Score method	2.18 (1.85-2.56)	1.80 (1.52-2.13)	1.74 (1.60-1.90)	1.49 (1.36-1.64)
Logistic regression	2.40 (2.07-2.79)	1.98 (1.68-2.33)	2.04 (1.87-2.22)	1.68 (1.53-1.84)
Factor analysis	2.21 (1.91-2.55)	1.81 (1.54-2.13)	1.61 (1.48-1.75)	1.35 (1.24-1.46)
Grade of membership	2.34 (1.89-2.89)	1.75 (1.40-2.19)	1.59 (1.46-1.73)	1.34 (1.22-1.47)

Note. Age, race, and poverty income ratio were adjusted for.

Table 3-7: The area under the ROC curve of allostatic load indices by the 5 scoring methods

	General Health		Diabetes		Hypertension	
	AUC	95% CI	AUC	95% CI	AUC	95% CI
Count-based method	0.73	0.68-0.79	0.83	0.79-0.87	0.74	0.72-0.77
Z-Score method	0.75	0.69-0.81	0.87	0.83-0.91	0.75	0.73-0.77
Logistic regression	0.75	0.69-0.81	0.92	0.88-0.95	0.79	0.77-0.81
Factor analysis	0.73	0.67-0.80	0.90	0.86-0.94	0.74	0.72-0.77
Grade of membership	0.72	0.66-0.78	0.82	0.78-0.86	0.70	0.67-0.73

Table 3-8: The area under the ROC curve of tailored allostatic load indices by the 5 scoring methods

	Diabetes		Hypertension	
	AUC	95% CI	AUC	95% CI
Count-based method	0.75	0.71-0.79	0.68	0.66-0.71
Z-Score method	0.76	0.72-0.81	0.68	0.66-0.71
Logistic regression	0.78	0.74-0.82	0.72	0.69-0.74
Factor analysis	0.78	0.73-0.82	0.70	0.67-0.73
Grade of membership	0.75	0.70-0.79	0.65	0.63-0.68

Table 3-9: Evaluations of the 5 scoring methods

Scoring Methods	Strengths	Weaknesses
The count-based method	<ul style="list-style-type: none"> • Simple; • Most frequently used; • Use natural units (i.e., number of indicators within high risk quartiles). 	<ul style="list-style-type: none"> • Discretizing variables loses information regarding the potential variability in their contribution in relation to overall risk; • Fails to consider the unequal weights of each indicator in the index.
The Z-Score method	<ul style="list-style-type: none"> • Simple; • The continuous function of biological measures accounts for available variance and increase statistical power. 	<ul style="list-style-type: none"> • Fails to consider the unequal weights of each indicator in the index; • More difficult interpretation due to standardization and loss of natural units.
Logistic regression	<ul style="list-style-type: none"> • Allows for unequal weights for each indicator. 	<ul style="list-style-type: none"> • Incorporates information on subsequent outcomes; • No prior AL studies have used it to assign weights to each indicator.
Factor analysis	<ul style="list-style-type: none"> • Allows for unequal weights for each indicator; • Does not incorporate information on subsequent outcomes. 	<ul style="list-style-type: none"> • The number of factors could be subjectively determined if not set at 1 a priori; • No prior AL studies have used it to assign weights to each indicator.
Grade of membership	<ul style="list-style-type: none"> • Allows for unequal weights for each indicator; • Does not incorporate information on subsequent outcomes. 	<ul style="list-style-type: none"> • The number of pure-type profiles is subjectively determined; • The method is challenging to produce.

Table 3-10: The cut-off points, sensitivities, and specificities of the allostatic load index by the count-based method for predicting general health, diabetes, and hypertension

Cut-off point as number of high risk quartiles	General Health		Diabetes		Hypertension	
	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
0	0.96	0.19	0.996	0.19	0.95	0.21
1	0.87	0.43	0.93	0.43	0.85	0.46
2	0.73	0.61	0.85	0.62	0.72	0.66
3	0.59	0.75	0.75	0.76	0.56	0.79
4	0.42	0.85	0.63	0.86	0.41	0.88
5	0.31	0.91	0.47	0.92	0.27	0.94
6	0.19	0.96	0.35	0.96	0.17	0.97
7	0.13	0.98	0.19	0.986	0.08	0.99
8	0.06	0.995	0.08	0.996	0.03	0.998
9	0.005	0.999	0.01	1	0.005	1
10	0	1	0	1	0	1

Note. ALI, allostatic load index.

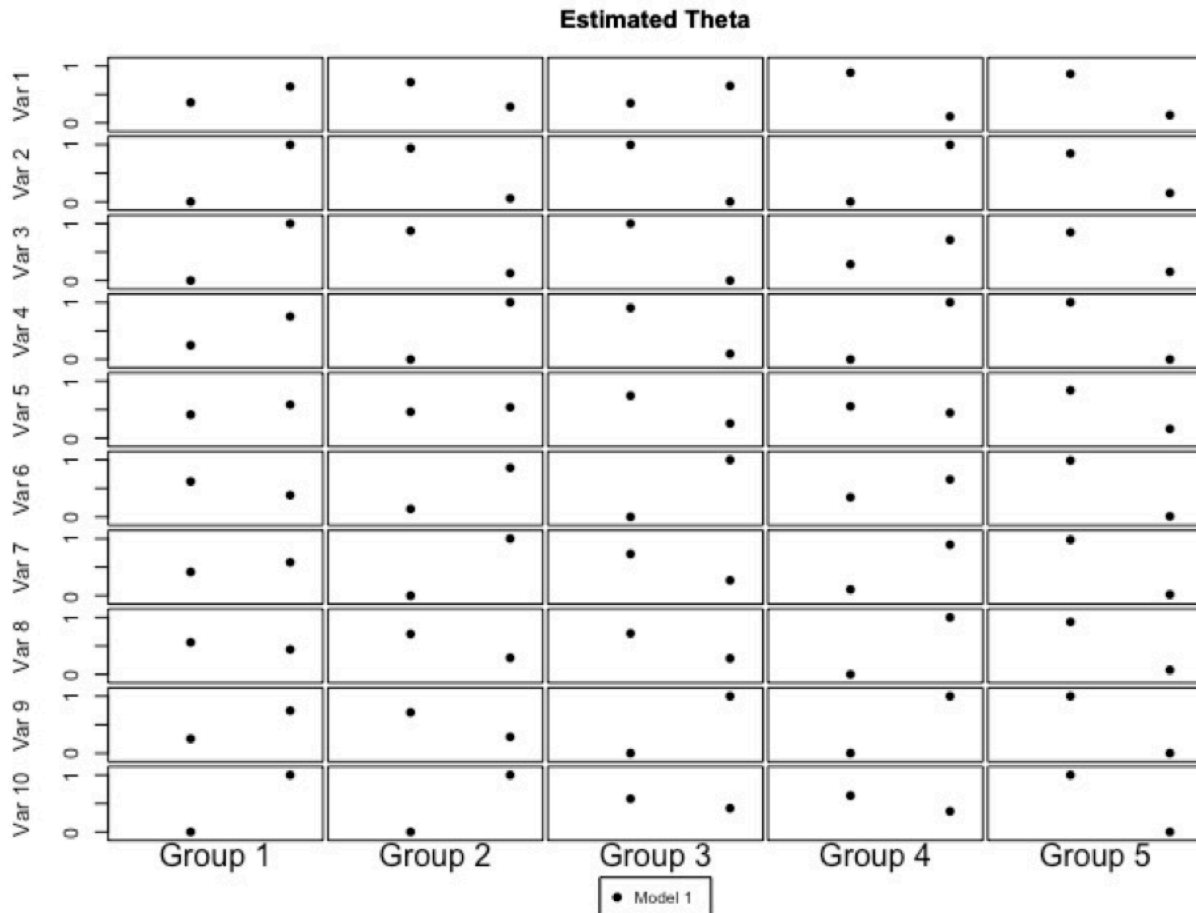


Figure 3-1A: The 5 pure-type profiles with the probability of each allostatic load indicator at low or high risk for poor health

Note. Variables from 1-10 are pulse, systolic blood pressure (SBP), diastolic blood pressure (DBP), body mass index (BMI), total cholesterol (TC), high-density lipoprotein (HDL), C-reactive protein (CRP), glycohemoglobin, glucose, and triglycerides.

The 5 pure types can be qualitatively labeled as follows: (1) high risk for hypertension (the probability of low levels of SBP and DBP equal to 0) and moderate risk for obesity and cardiovascular diseases (the probability of low levels of BMI and TC less than 0.5 and triglycerides equal to 0); (2) high risk for obesity and cardiovascular diseases (the probability of low levels of BMI and triglycerides equal to 0 and high levels of HDL less than 0.5), but low risk for hypertension and diabetes (the probability of low levels of SBP, DBP, glucose, and glycohemoglobin greater than 0.5); (3) high risk for diabetes and cardiovascular diseases (the probability of high levels of HDL and low levels of glucose equal to 0), but not at risk for hypertension and low risk for obesity (the probability of low levels of SBP and DBP equal to 1 and BMI greater than 0.5); (4) high risk for hypertension, obesity, diabetes and moderate risk for cardiovascular diseases (the probability of low levels of BMI, glucose, glycohemoglobin, and SBP equal to 0 and low levels of DBP and high levels of HDL less than 0.5); (5) the lowest risk for poor health outcomes (the probability of low levels of all indicators greater than 0.8).

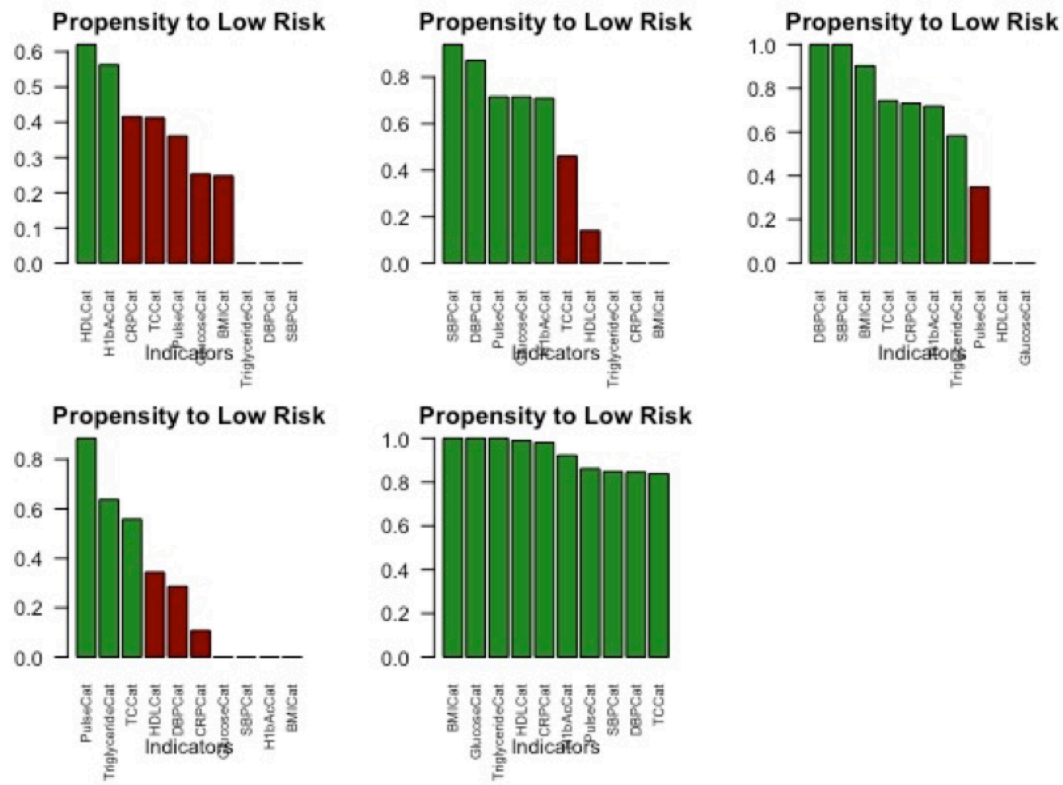
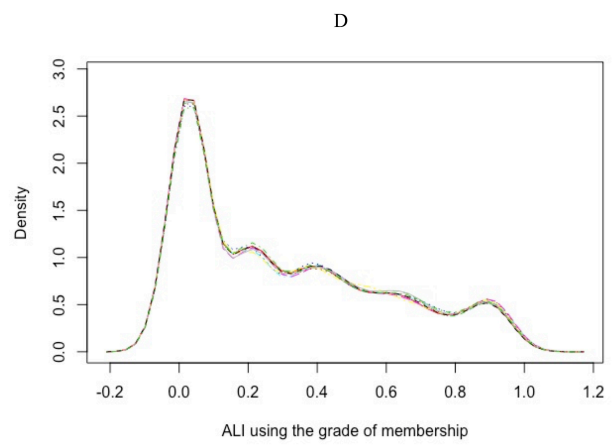
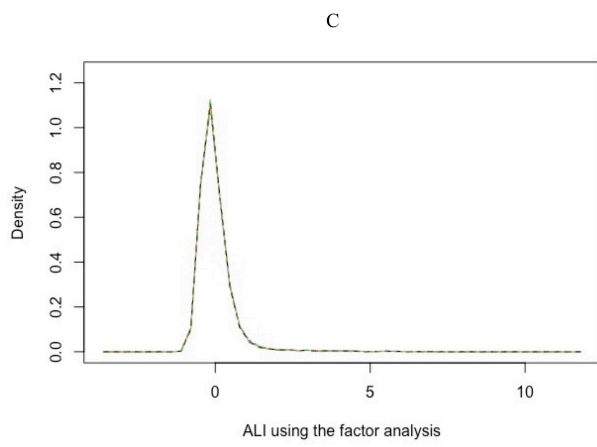
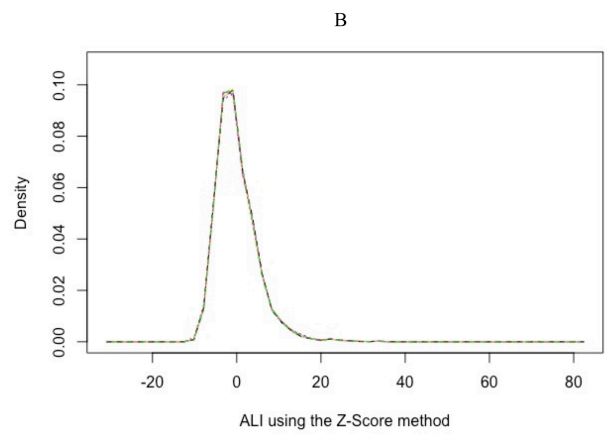
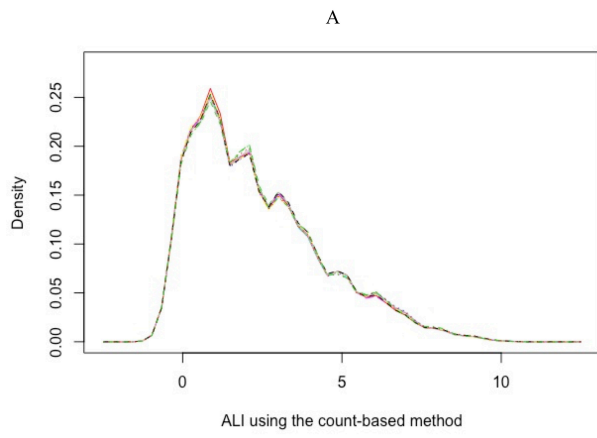


Figure 3-1B: The probability of each allostatic load indicator at low risk for poor health in the 5 pure-type profiles

Note. From top left to right are Group 1, 2, and 3 and from bottom left to right are Group 4 and 5. The green bar indicates the probability of the indicator at low risk for poor health is higher than 0.5 and the red bar indicates the probability is 0.5 or lower.

The 5 pure types can be qualitatively labeled as follows: (1) high risk for hypertension (the probability of low levels of SBP and DBP equal to 0) and moderate risk for obesity and cardiovascular diseases (the probability of low levels of BMI and TC less than 0.5 and triglycerides equal to 0); (2) high risk for obesity and cardiovascular diseases (the probability of low levels of BMI and triglycerides equal to 0 and high levels of HDL less than 0.5), but low risk for hypertension and diabetes (the probability of low levels of SBP, DBP, glucose, and glycohemoglobin greater than 0.5); (3) high risk for diabetes and cardiovascular diseases (the probability of high levels of HDL and low levels of glucose equal to 0), but not at risk for hypertension and low risk for obesity (the probability of low levels of SBP and DBP equal to 1 and BMI greater than 0.5); (4) high risk for hypertension, obesity, diabetes and moderate risk for cardiovascular diseases (the probability of low levels of BMI, glucose, glycohemoglobin, and SBP equal to 0 and low levels of DBP and high levels of HDL less than 0.5); (5) the lowest risk for poor health outcomes (the probability of low levels of all indicators greater than 0.8). SBP, systolic blood pressure; DBP, diastolic blood pressure; BMI, body mass index; TC, total cholesterol; HDL, high-density lipoprotein.



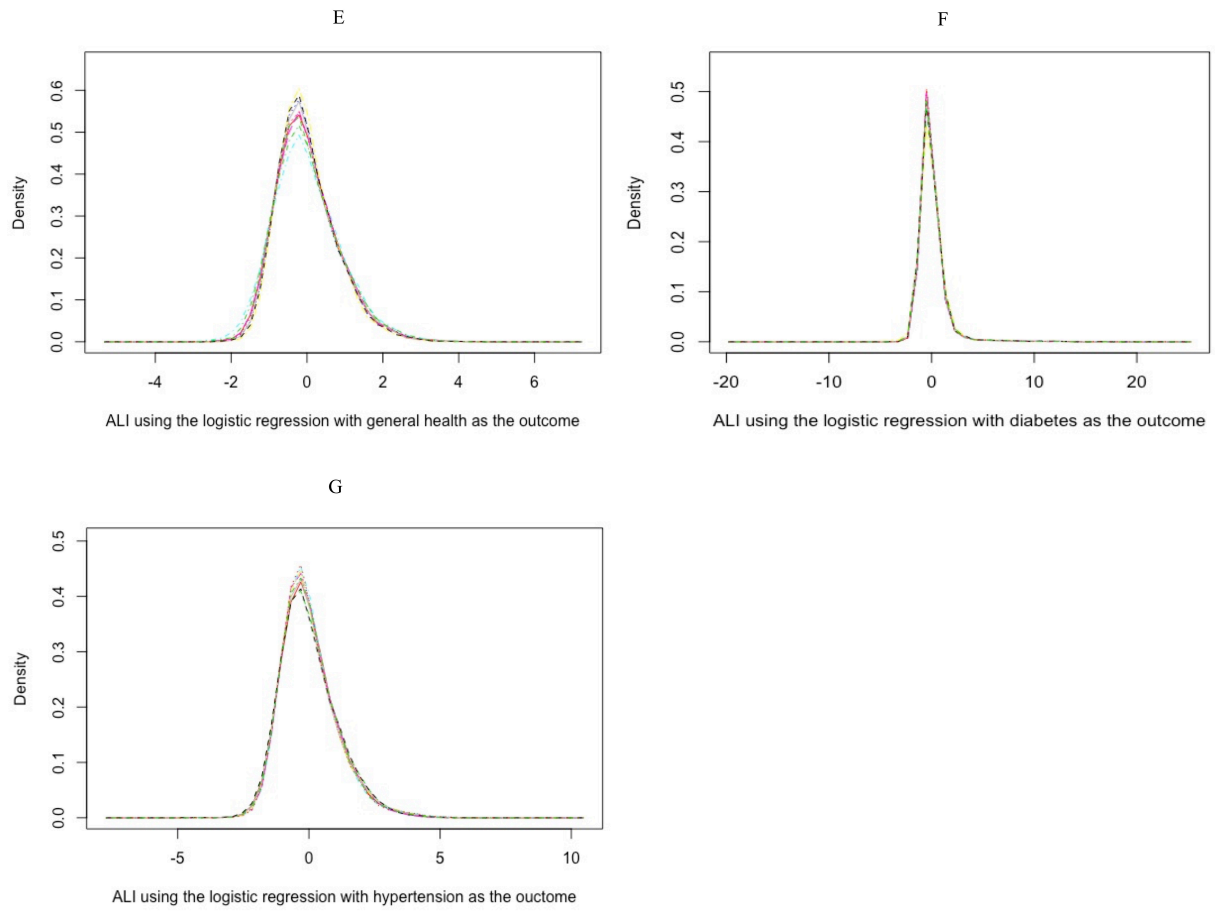


Figure 3-2: The distributions of allostatic load indices using the 5 scoring methods

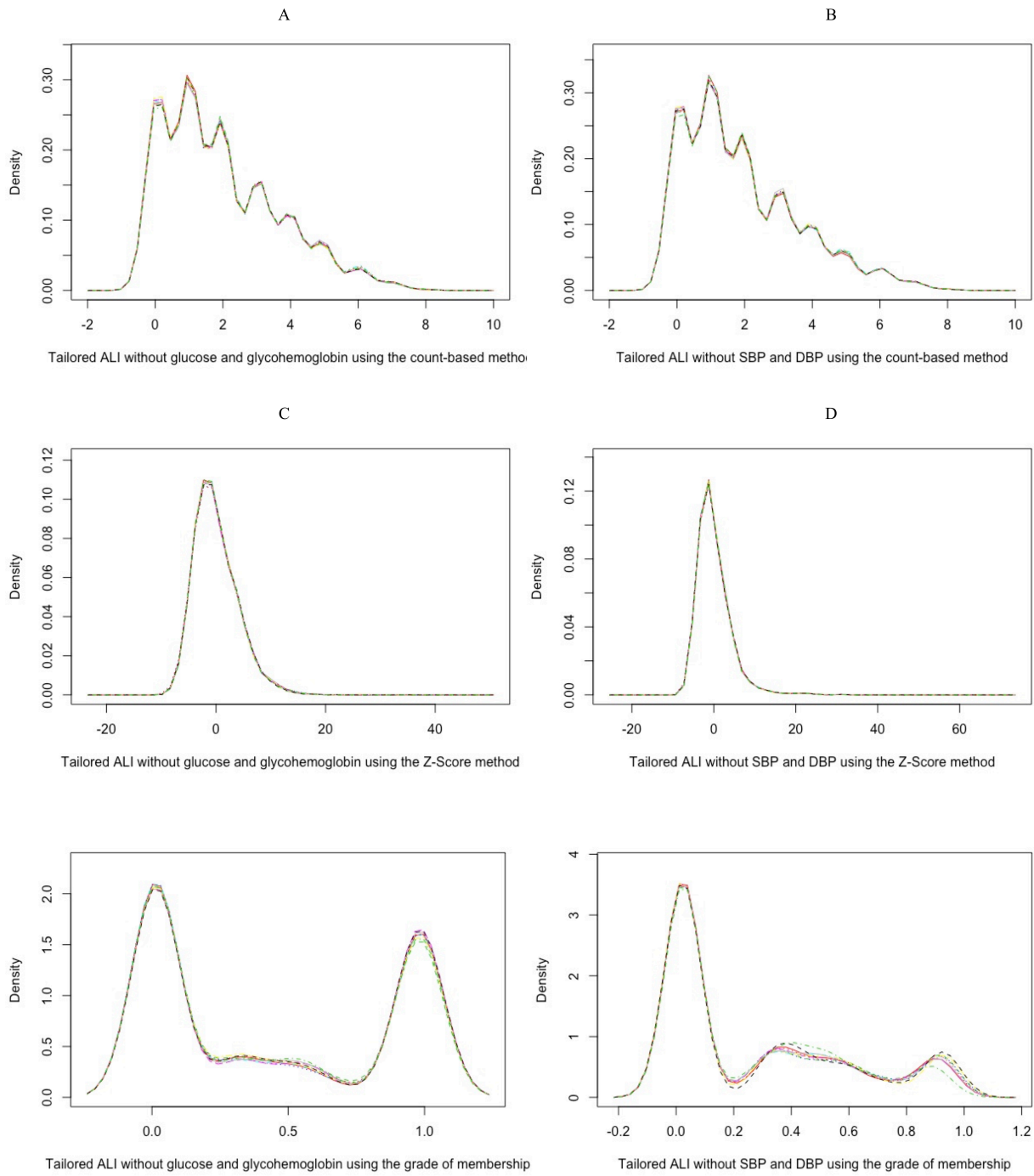


Figure 3-3: The distributions of tailored allostatic load indices using the 5 scoring methods

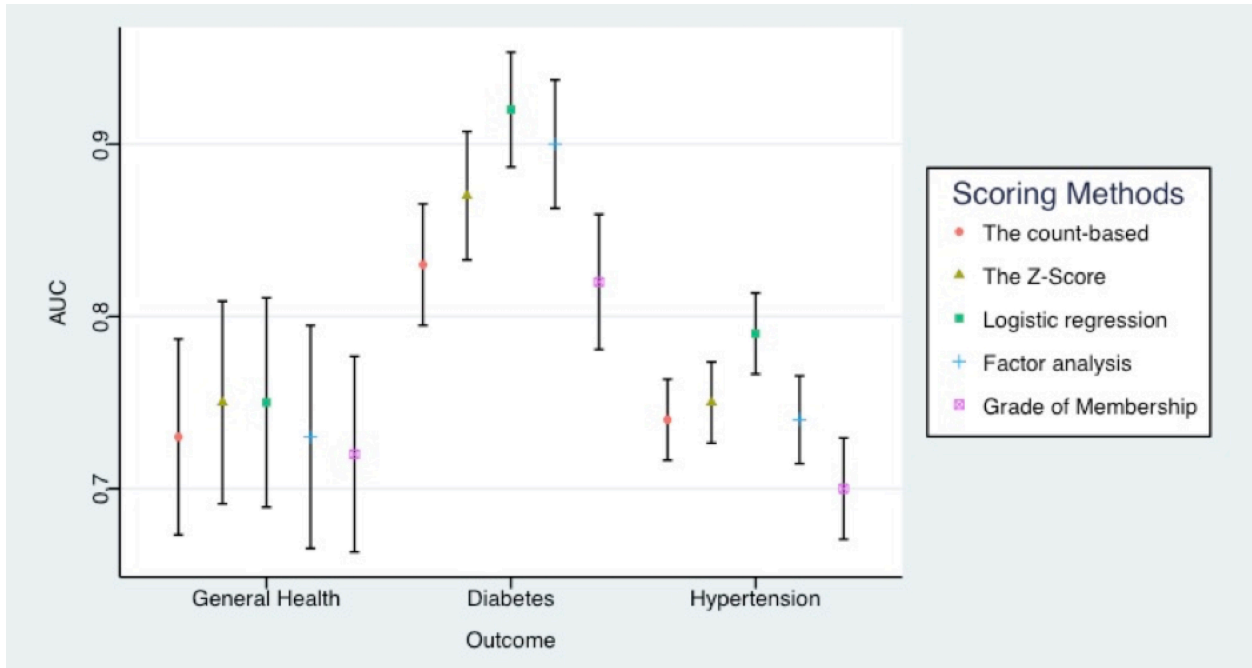


Figure 3-4: The ROC area statistics and 95% confidence intervals of each allostatic load summary measure in terms of general health, diabetes, and hypertension

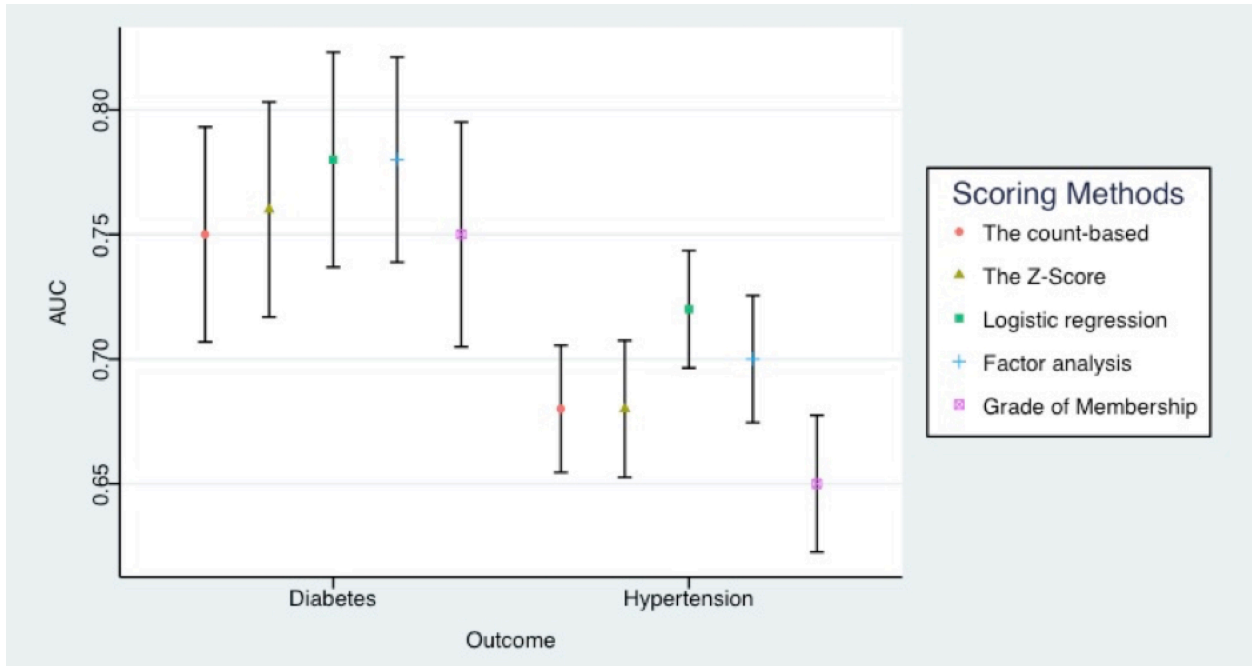


Figure 3-5: The ROC area statistics and 95% confidence intervals of each tailored allostatic load summary measure in terms of diabetes and hypertension

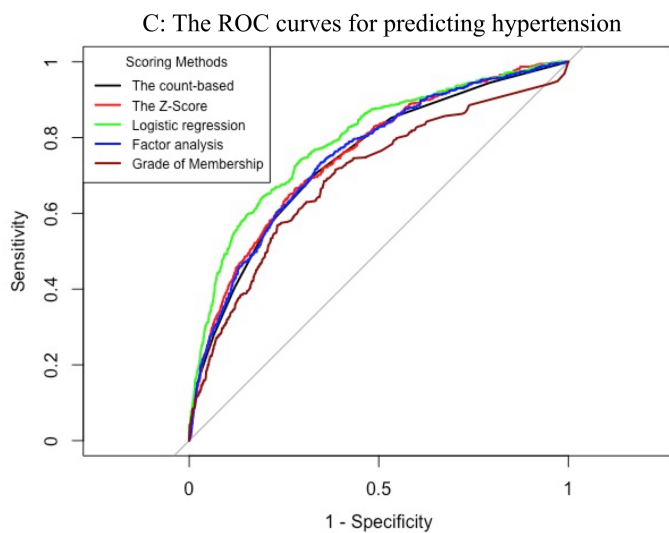
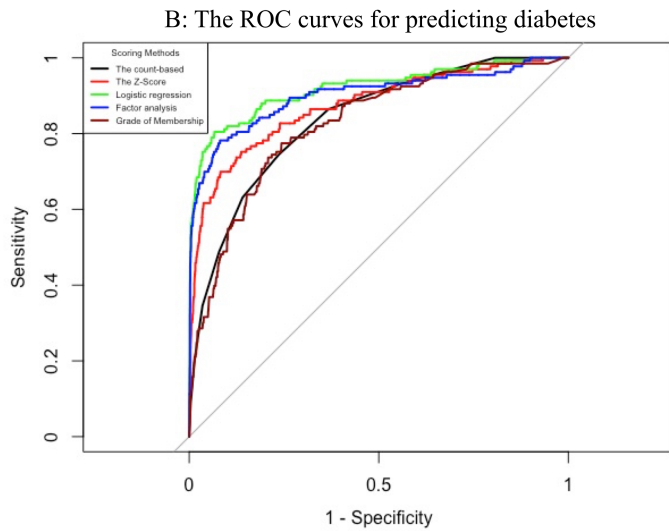
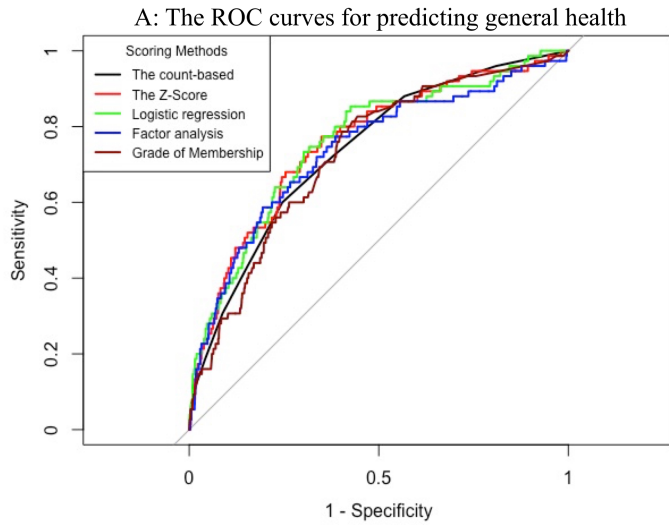
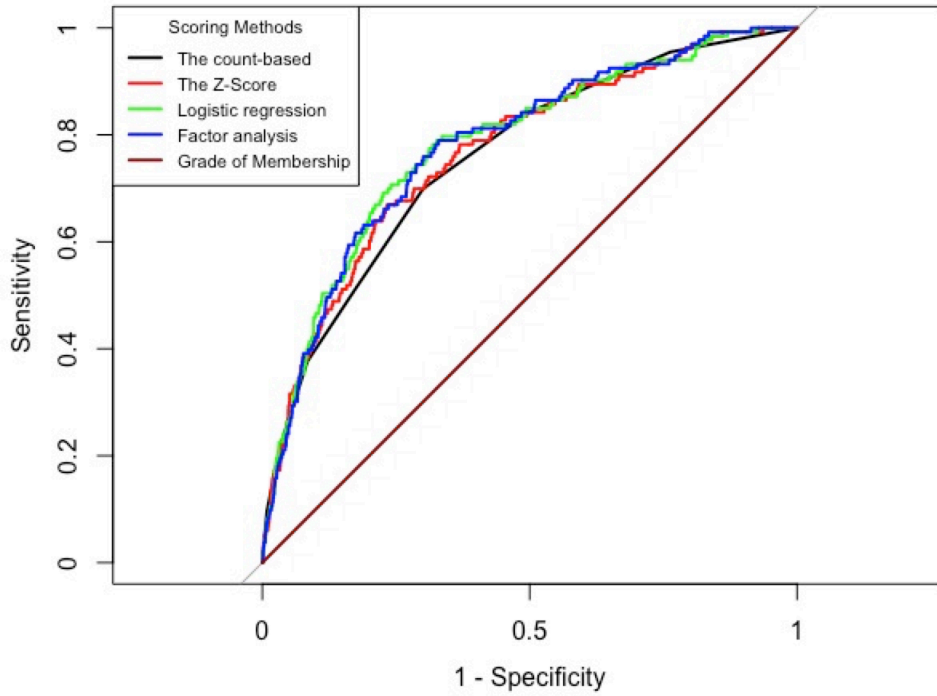


Figure 3-6: The ROC curves of allostatic load indices by the 5 scoring methods for predicting general health, diabetes, and hypertension

A: The ROC curves for predicting diabetes



B: The ROC curves for predicting hypertension

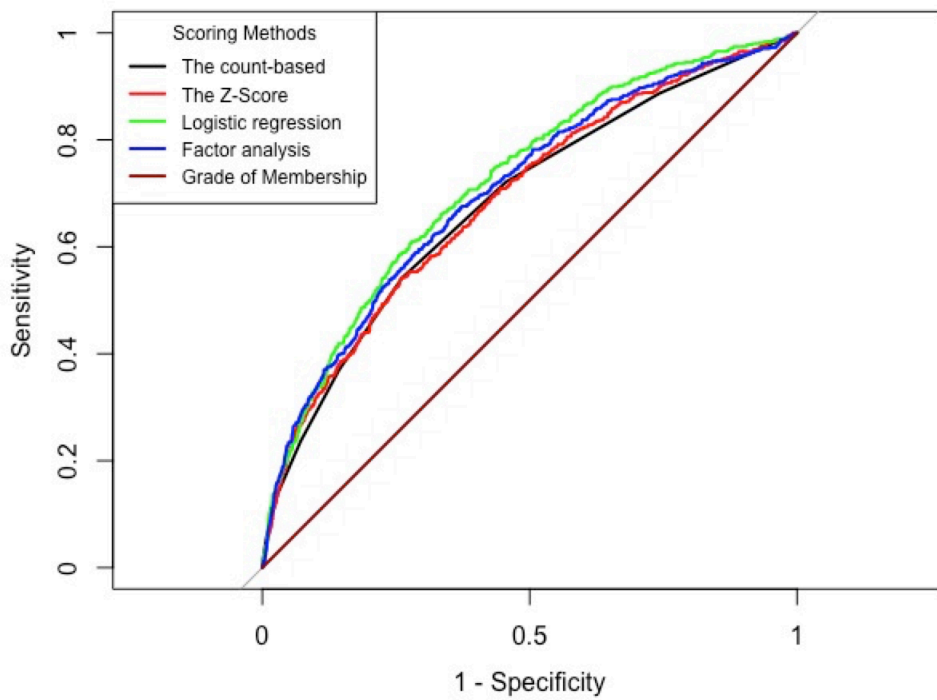


Figure 3-7: The ROC curves of tailored allostatic load indices by the 5 scoring methods for predicting diabetes and hypertension

Chapter 4: Exploring the validity of allostatic load in pregnant women

Abstract

The allostatic load theory can be applied to perinatal outcomes research to understand the biological pathway for the impact of maternal chronic stress on adverse perinatal outcomes. However, there have been few pregnancy allostatic load studies. Due to physiological changes in pregnancy, whether allostatic load in pregnancy could validly reflect women's long-term physiological functions is unclear. The predictive validity of pregnancy allostatic load index on adverse birth outcomes is also unknown. This study aimed to test whether the allostatic load index has face validity in pregnancy and to assess the predictive validity of the index on a proxy for adverse birth outcomes, having had a premature low birth weight infant on the previous pregnancy. This is a secondary analysis of data on pregnant women from the 1999-2006 National Health and Nutrition Examination Survey (NHANES). We found the allostatic load summary score remained steady across gestational months and was not statistically different from the average score in the non-pregnant population of the NHANES. We also found elevated allostatic load index scores were associated with history of adverse birth outcomes. Our study suggested that measuring AL in pregnancy could reflect women's true physiological functions when gestational age was considered when scoring AL. However, the score could only identify 31% of women who had a prior adverse birth outcome, a predictive performance similar to other obstetric risk assessments, but unsatisfactory. Combining the allostatic load summary score,

existing risk scoring systems, and technical assessments may improve the accuracy for predicting preterm birth.

Keywords: allostatic load, pregnant women, validity, preterm birth, low birth weight

Introduction

Adverse perinatal outcomes such as preterm birth and low birth weight are significant public health concerns. In 2016, around 1 in 10 were born before 37 weeks of gestation in the United States (Martin, Hamilton, & Osterman, 2017). According to the 2015 data from the National Vital Statistics System, preterm birth and low birth weight account for approximately 17% of infant mortality (Xu, Murphy, Kochanek, & Arias, 2016). In 2005, the Institute of Medicine reported that the annual cost associated with preterm birth in the United States was more than 26 billion (Butler, & Behrman, 2007). A growing body of perinatal research has demonstrated maternal psychosocial and traumatic stress (i.e., child maltreatment, low socioeconomic status) could contribute to adverse perinatal outcomes (Hellgren, Akerud, Skalkidou, & Sundstrom-Poromaa, 2013; King et al., 2010; Shaw et al., 2014; Shea et al., 2007; Voegtline et al., 2013). However, the biological mechanisms that may mediate the effects of maternal psychosocial and traumatic stress on perinatal outcomes remain to be elucidated.

Allostatic load (AL) refers to the cumulative dysregulations of multiple physiological systems responsive to chronic stress. It has been well documented in relation to both chronic stress and adverse health outcomes in non-pregnant populations (Beckie, 2012). Some studies suggested AL played a mediating role in the link between stress and health outcomes (Crimmins, Kim, & Seeman, 2009; Seeman et al., 2004). Applying the AL concept to perinatal research, it may serve as a potential contributor to adverse perinatal outcomes in women who experienced psychosocial and traumatic stress.

However, pregnancy is a state that involves temporary alterations in function of physiology across multiple body systems to facilitate reproduction. Pregnancy physiology disrupts the normal function of multiple physiological systems. During pregnancy, physiological stress systems that constitute AL are modified and even more complicated than in the non-pregnant state (Blackburn, 2014). The physiological changes that result from maternal stress may differ from those of the non-pregnant state. A review of 15 studies that used standardized laboratory stressors to test stress reactivity to pain and psychological stress during pregnancy found that physiological stress responses (i.e., blood pressure, heart rate, cortisol, epinephrine, norepinephrine) to exogenous challenges might be attenuated during pregnancy (de Weerth & Buitelaar, 2005). Few studies have measured AL during pregnancy because they may have a concern that the changing levels of AL-related indicators during pregnancy may not reflect women's true AL. It is a question of the extent to which the great changes in physiology and physiological stress reactivity during pregnancy affect AL indicators from neuroendocrine, immune, cardiovascular, and metabolic systems in ways that reduce the validity of AL measurement. It is unknown whether AL in pregnancy could validly reflect women's long-term physiological functions. Although our previous study found satisfactory predictive validities of AL index (ALI) by different scoring methods on self-reported general health status, diabetes, and hypertension in women of reproductive age, the predictive validity of pregnancy ALI on adverse birth outcomes is unknown.

This study used the National Health and Nutrition Examination Survey (NHANES) data from 1999-2006 to address two aims: A) To examine the gestational patterns of AL indicators and the ALI and to test whether the ALI has face validity in pregnancy despite physiological changes of pregnancy; and B) To assess the predictive validity of pregnancy ALI on a proxy for

adverse birth outcomes, having had a premature low birth weight infant on the previous pregnancy. We followed the STROBE checklist to report the study in the paper (von Elm et al., 2008).

Methods

Study design

This is a secondary analysis of data from the NHANES that is a cross-sectional study with a large nationally representative sample. It consists of the survey administered in the home and the physical examination in a mobile examination center (MEC). To address the study aims, we used data from the 1999 to 2006 cycles of the NHANES. The data collected after 2006 were not used because since 2007 the pregnant sample has not been oversampled, due to disclosure risks only data on females between 20 and 44 who had urine pregnancy results has been available for public use since 2007, as well as no data on adverse perinatal outcomes have been collected. This secondary analysis of data was exempt from IRB review because it was done via the de-identified dataset.

Because of the cross-sectional study design of the NHANES, information on perinatal outcomes for pregnant participants is unknown. However, the NHANES queried past history of delivering low birth weight and premature infants. Women with a history of preterm delivery had an around 3-fold increased risk of preterm delivery in a subsequent pregnancy and women who delivered a very low birth weight first infant experienced an around 12-fold increased risk of delivering a low birth weight second infant (Bratton, Shoultz, & Williams, 1996; Mazaki-Tovi et al., 2007). Thus, to address Aim B, we considered history of delivering a premature low birth weight infant as a proxy/high risk for adverse perinatal outcomes of the current pregnancy.

Setting

The health interview, including sociodemographic information, is administered in the participant's home. The physical examination and computer-assisted personal interview on sociodemographic and outcome variables took place in MECs that traveled to locations throughout the country. Details are presented in the National Health and Nutrition Examination Survey: Plan and Operations, 1999–2010 (Zipf et al., 2013).

Participants

Pregnant women were included in the study. Pregnancy status at the time of the MEC examination was reported for females 8-59 years of age through the urine pregnancy test. If the urine test was positive, the status is coded as being pregnant at examination. A total of 1256 women were pregnant at the exam.

To address Aim A (the descriptive analyses), a total of 1056 pregnant women were included. One hundred ninety-eight pregnant women (15.8%) did not report their gestational month at the exam and 2 women (0.2%) reported they were at gestational month 10. These 200 women were excluded.

To address Aim B (the predictive validation), the sample is smaller (N=665) because we need to include only those who also had a previous birth. Women, who were pregnant at the exam and have been pregnant at least twice, including the current pregnancy, were included for analyses so that the outcome (history of adverse birth outcome) is available on the previous pregnancy. Women who were pregnant at the exam for the first time were excluded. Women who had missing data on the key outcome variable (history of adverse birth outcomes) were also excluded. Among 1256 pregnant women, 330 (26.3%) were pregnant at exam for the first time and 110 (8.8%) did not report how many times they have been pregnant. Among the remaining

816 pregnant women, 79 (6.3%) had missing data on the outcome variable and 60 (4.8%) did not report what month of pregnancy they were in. We also excluded 2 women (0.2%) reporting they were at gestational month 10 and 10 women (0.8%) whose children were low birth weight but not born before 37 weeks of gestation. Thus, 665 pregnant participants were included for analysis to address Aim B.

Variables and data sources

Detailed information for measures was in Table 4-1.

AL scoring

In previous work we demonstrated that the logistic regression method of scoring the ALI is optimal in research situations when the outcome information is available, otherwise the simpler count-based method is a good alternative with good predictive validity for predicting unknown outcomes as is the case in clinical use. For the descriptive analysis (Aim A; $N=1056$), we used the count-based method to score the ALI because the outcome information (history of adverse birth outcomes) is not applicable. For the predictive validation (Aim B; $N=665$), we included prior adverse birth outcomes as the outcome variable, but to be consistent through the study, we still used the count-based method to score the ALI.

Because of the cross-sectional study design of the NHANES, pregnant women were at different gestational month. Furthermore, we expected the levels of each indicator to change across pregnancy. Thus, the ALI calculation by the count-based method accounts for gestational month. We divided the whole sample into 9 gestational-age-specific subsamples in which participants were at the same gestational month (gestational month ranges 1-9). The ALI by the count-based method was the number of indicators on which participants scored in the top risk quartile (except for HDL in the bottom quartile) of each gestational-age-specific subsample's

distribution, yielding a possible score range of 0 to 10. That is, we calculated the risk quartile for each indicator based on the distribution of each of the 9 subsamples in which participants were at the same gestational month.

Bias

NHANES is an ongoing, nationally representative study. Data are collected with standardized procedures and protocols developed and validated by the National Center for Health Statistics (NCHS) for all household interview, clinical examinations and laboratory tests. This helps to assure that the data for this analysis is of high quality in terms of reliability and validity.

Statistical methods

Preliminary analyses. Means, standard deviations, frequencies, and percentages were used to describe sociodemographic characteristics, gestational month, and history of adverse birth outcomes.

Handling of missing data. “Refused” or “don’t know” responses were recoded as missing values. In the sample of 1056 pregnant women, 13.9% of data was missing. In the sample of 665 pregnant women, 12.5% of data was missing. Since the percentages of missing data were less than 15% in both samples, we did not impute missing data in analyses.

Aim A analyses (the descriptive analysis). The averages of the ALI score and each AL indicator were plotted for each gestational-month group after adjusting for age, race, and poverty income ratio to describe changes in the ALI score and each individual indicator across different gestational month. The average ALI score by the count-based method in the non-pregnant population from the NHANES 2001-2006 dataset was added in the plot to compare with the averages of pregnancy ALI scores. The ALI scores in women with history of adverse birth outcomes were also added in the plot. Linear regression models were conducted to examine the

effects of gestational month on each indicator levels. Each indicator was included respectively as the dependent variable and age, race, and poverty income ratio were adjusted for in the regression models. Based on the plots, gestational month was included as the independent variable when the relationships between gestational month and indicators were linear, while trimester was included as the independent variable when the relationships were not linear.

Aim B analyses (the predictive validation). Binomial logistic regression models with the ALI score as the independent variable and history of adverse birth outcomes as the dependent variable were conducted to examine the association between the ALI score and history of adverse birth outcomes. Age, race, and poverty income ratio were also adjusted for. Using the count-based method, cut-off points, sensitivities, specificities, positive predictive value (PPV), and negative predictive value (NPV) were also computed. All statistical analyses were performed using R Software Version 3.4.2 (R Core Team, 2017).

Results

The sample characteristics

In the sample ($N=1056$) to address Aim A, the mean age of the sample was 27 and 13.7% were non-Hispanic black. The average poverty income ratio was 2.4. Around 32% of the sample had less than high school education and nearly 77% were married or living with partner. In the sample ($N=665$) to address Aim B, the mean age of the sample was approximately 28 and 13.7% were non-Hispanic black. The average poverty income ratio was 2.1, 34% of the sample had less than high school education, and 80% were married or living with partner. Thirty-one women (4.7%) reported history of adverse birth outcomes. The sample characteristics were shown in Table 4-2.

The main findings for Aim A

As seen in Figure 4-1, the ALI score remained steady across gestational month and the scores in each gestational month were around the average ALI score in the non-pregnant population from our previous study ($M=2.35$, $SE=0.03$, $N=4319$). The average ALI score in the non-pregnant population was within the one standard error above and below the average ALI score in each gestational month. The ALI score in pregnant women was not significantly different from the average ALI score in the non-pregnant population. In addition, among pregnant women with prior adverse birth outcomes, more women had ALI scores above the average score in the pregnant sample. The summary of the 10 indicators in terms of the changing patterns in pregnancy was presented in Table 4-3 and the changing patterns across gestation were shown in Figure 4-2.

The main findings for Aim B

As seen in Table 4-4, the binomial logistic regression showed that women with a history of adverse birth outcomes had elevated ALI scores compared to women without a history of adverse birth outcomes ($OR=1.24$, $95\%CI=1.002-1.52$). The association remained significant after adjusting for age, race, and poverty income ratio ($OR=1.25$, $95\%CI=1.002-1.56$). Using the count-based method, when the cut-off point of the ALI score is 3, the sensitivity was 31% and the specificity was 79% (Table 4-5); and the PPV was 0.067 and the NPV was 0.96.

Discussion

The study described changes in the ALI score and each individual AL indicator across different gestational month as well as assessed the predictive validity of pregnancy AL for predicting prior adverse birth outcomes as a proxy for adverse birth outcomes of the current pregnancy. We found significant differences in levels of each AL indicator at different

gestational month except for C-reactive protein (CRP). The ALI score remained steady across gestational month and the summary score at each gestational month was not different from the average score in the non-pregnant population of the NHANES. We also found higher ALI score was associated with history of adverse birth outcomes. But the predictive validity of the ALI score was not satisfactory.

AL indicators change across gestational month. Pulse, total cholesterol (TC), and triglycerides increased with the progression of pregnancy. Systolic blood pressure (SBP), diastolic blood pressure (DBP), body mass index (BMI), glycohemoglobin, and glucose decreased up to the middle of gestation and then increased up to the late gestation, with the peak level reach at the second trimester. Regarding high-density lipoprotein (HDL) and CRP, there was an increase from early gestation to the middle of gestation and then a decrease in the late gestation. The changing patterns of each indicator across gestation were consistent with known pregnancy physiology. Except for CRP, there were significant differences in each AL indicator levels across different gestational month. This suggests that gestational age should be taken into account when scoring AL in pregnant women. Additionally, population-based cut-off points need to be determined for each AL indicator at each gestational month. Given that the NHANES has a nationally representative sample with data collected with standardized procedures and protocols, future research could identify these cut-off points using the NHANES data.

The ALI score at each gestational month was not different from the average AL summary score in the non-pregnant population. This indicates that measuring AL at any time in pregnancy would reflect women's true physiological functions. But the influence of gestational age should be considered when scoring AL.

We found that women with history of adverse birth outcomes had higher pregnancy ALI score than those without history of adverse birth outcomes. Our finding is consistent with a study conducted in New Orleans that reported negative associations of pregnancy AL with gestational age in 42 women (Wallace et al., 2013). A study also used the NHANES data but focused on non-pregnant women, which found higher non-pregnancy AL levels in women with history of small for gestational age or preterm birth compared to those with normal birth outcomes (Hux, Catov, & Roberts, 2014). The findings from the study by Hux et al. (2014) and our study suggest that women with history of adverse birth outcomes may have elevated AL levels either in a subsequent pregnancy or in a non-pregnancy state. Those high-risk women may already have had elevated AL levels before their first birth.

In obstetric practice, history of a prior adverse birth outcome is assessed as an important risk factor for a subsequent adverse birth outcome. The ALI score by the count-based method was associated with prior adverse birth outcomes, but not at a level sufficiently accurate to serve as a stand-alone measure of risk to predict history of future adverse birth outcomes. Using the cut-off point of 3, the ALI can only identify 31% of women who actually reported history of adverse birth outcomes, while a substantial proportion of women reporting history of adverse birth outcomes would be classified as low risk. The ALI was reasonably specific, but not adequately sensitive to use as a sole predictor of risk for adverse birth outcomes. We tested the predictive performance of the ALI for predicting prior adverse birth outcomes rather than the outcomes subsequent to pregnancy, which is a significant limitation of our study. However, use of the NHANES sample, which is representative of the US population, allows us to describe cut-off points that are tentatively valid for clinical use. Due to the limitation of the NHANES data, we might be able to recommend a cut-off point of pregnancy ALI after research on data where

the outcome is adverse birth outcomes subsequent to the current pregnancy instead of prior adverse birth outcomes.

The poor predictive performance of the AL summary score is similar to the performances of existing risk-scoring systems and technical screening assessments. A review study synthesized the validity of the most widely used Creasy risk-scoring system and found most studies reported a low sensitivity of around 25%-45% and a specificity of over 80% in predicting preterm birth on different populations (Edenfield, Thomas, Thompson, & Marcotte, 1995). Another review study found a wide range of accuracy of 12 published risk-scoring systems in predicting preterm birth and suggested that no existing systems of scoring could be recommended for clinical practice (Honest et al., 2004). A study systematically reviewed published studies to examine the accuracy of the combination of fetal fibronectin (fFN) and ultrasound examination of cervical length (CL) as screening tools for preterm birth in symptomatic patients. The combination of fFN and CL had a sensitivity of 36.8% and a specificity of 83.0% for predicting preterm birth at less than 37 weeks of gestation (DeFranco, Lewis, & Odibo, 2013). The insufficient accuracies of the ALI score, existing risk-factor screening, and technical screening tools probably result from the multifactorial nature of preterm birth. According to the 2005 research agenda for preterm birth, genetic, environmental, psychosocial, and behavioral factors interact in complex pathogenesis and multiple biological pathways leading to preterm birth (Green et al., 2005). An effective screening approach for preterm birth may need to include multiple factors that predict preterm birth, including sociodemographic and stress characteristics, reproductive history, medical conditions, findings on a physical examination, the fFN test result, the CL by ultrasound measurement, or measurement of physiological indicators. Goldenberg et al. (2005) suggested combining tests of multiple biomarkers with fFN and CL tests might enhance the predictive

accuracy for preterm birth (Goldenberg, Goepfert, & Ramsey, 2005). Therefore, the use of a combination of the ALI score with existing risk scoring systems and technical assessments may improve the overall accuracy compared with single approach alone. Future research needs to explore whether a comprehensive screening approach would have a high predictive performance for preterm birth.

In contrast to the sensitivity and PPV, we found high specificity and NPV with the cut-off point of 3. The AL summary measure may identify pregnant women who are likely to be at low risk for adverse birth outcomes. For those women, intensive and early surveillance and treatment can be avoided. The distress caused by the risk diagnosis can be also avoid. Furthermore, unnecessary medical costs, resources, and human efforts can be reduced.

This study has limitations. Data on biomarkers from the neuroendocrine system as well as perinatal outcomes (i.e., preterm birth and low birth weight) are lacking in the NHANES database. The ALI was constructed without biomarkers from the neuroendocrine system and prior adverse birth outcomes were included as the outcome variable, which may affect the predictive performance of the ALI. A prospective study using a full complement of physiological indicators and following up women's perinatal outcomes is needed to replicate our analyses. The study has strengths as well. The NHANES database includes numerous physiological measures from multiple body systems and a substantial amount of pregnancy physiological data, which provides adequate AL indicators for our study to create pregnancy ALI. Additionally, the NHANES data are collected with standardized procedures and protocols to assure that the data for this analysis is of high quality in terms of reliability and validity.

This study is an initial step and more work is definitely needed to test whether AL can be a sensitive measure for screening pregnant women at risk for adverse birth outcomes. But, this

study may still have a few potential implications. First, this study filled the gap regarding how to score AL in pregnancy. For future studies that recruit women at different gestational age, gestational age should be taken into account when scoring AL. Second, the study of these multi-system physiological indicators and applying the AL theory to perinatal outcomes research may provide important insights into the potential biological mechanisms or pathways ultimately leading to preterm birth. Third, given the high specificity and NPV of the AL summary score, it may have potential use in identifying pregnant women who are likely to be at low risk for adverse birth outcomes. Thus, unnecessary prenatal care and treatment can be avoided. The last but not the least, the ALI score as a new risk assessment method may contribute to developing a comprehensive assessment of risk for preterm birth. The AL summary measure may be added to existing risk scoring systems and technical assessments to identify the majority of pregnant women subsequently having an adverse birth outcome. Whether offering antenatal care to asymptomatic pregnant women depends on highly accurate risk assessment. A valid systematic assessment for the risk of preterm birth could identify high-risk women, followed by ongoing surveillance and intensive, specialized care that aimed at preventing or delaying preterm birth. For pregnant women whose ALI score was particularly high, stress mitigation interventions in particular might prove valuable.

In summary, measuring AL in pregnancy can reflect women's true physiological functions, but gestational age needs to be taken into account when scoring AL. Although the predictive performance of the ALI score is not satisfactory, the ALI score, existing risk scoring systems, and technical assessments may perform better for predicting preterm birth when used in combination than used alone. Future studies need to test the predictive accuracy of the

combination of those scoring and testing. With a highly accurate risk assessment approach, high-risk women could be identified and referred to appropriate levels of antenatal care.

References

- Butler, A. S., & Behrman, R. E. (Eds.). (2007). *Preterm birth: causes, consequences, and prevention*. National Academies Press.
- Beckie, T. M. (2012). A systematic review of allostatic load, health, and health disparities. *Biological Research for Nursing, 14*(4), 311-346.
- Blackburn, S. (2014). *Maternal, fetal, & neonatal physiology*: Elsevier Health Sciences.
- Bratton, S. L., Shoultz, D. A., & Williams, M. A. (1996). Recurrence risk of low birthweight deliveries among women with a prior very low birthweight delivery. *American Journal of Perinatology, 13*(3), 147-150. doi: Doi 10.1055/S-2007-994313
- Crimmins, E. M., Kim, J. K., & Seeman, T. E. (2009). Poverty and Biological Risk: The Earlier "Aging" of the Poor. *Journals of Gerontology Series a-Biological Sciences and Medical Sciences, 64*(2), 286-292. doi: 10.1093/gerona/gln010
- de Weerth, C., & Buitelaar, J. K. (2005). Physiological stress reactivity in human pregnancy--a review. *Neuroscience & Biobehavioral Reviews, 29*(2), 295-312. doi: 10.1016/j.neubiorev.2004.10.005
- DeFranco, E. A., Lewis, D. F., & Odibo, A. O. (2013). Improving the screening accuracy for preterm labor: is the combination of fetal fibronectin and cervical length in symptomatic patients a useful predictor of preterm birth? A systematic review. *American Journal of Obstetrics and Gynecology, 208*(3). doi: ARTN 233.e110.1016/j.ajog.2012.12.015
- Edenfield, S. M., Thomas, S. D., Thompson, W. O., & Marcotte, J. J. (1995). Validity of the Creasy risk appraisal instrument for prediction of preterm labor. *Nursing Research, 44*(2), 76-81.
- Goldenberg, R. L., Goepfert, A. R., & Ramsey, P. S. (2005). Biochemical markers for the prediction of preterm birth. *American Journal of Obstetrics & Gynecology, 192*(5 Suppl), S36-46. doi: 10.1016/j.ajog.2005.02.015
- Green, N. S., Damus, K., Simpson, J. L., Iams, J., Reece, E. A., Hobel, C. J., . . . March Of Dimes Scientific Advisory Committee On, P. (2005). Research agenda for preterm birth: recommendations from the March of Dimes. *American Journal of Obstetrics & Gynecology, 193*(3 Pt 1), 626-635. doi: 10.1016/j.ajog.2005.02.106
- Hellgren, C., Akerud, H., Skalkidou, A., & Sundstrom-Poromaa, I. (2013). Cortisol awakening response in late pregnancy in women with previous or ongoing depression. *Psychoneuroendocrinology, 38*(12), 3150-3154. doi: 10.1016/j.psyneuen.2013.08.007
- Honest, H., Bachmann, L. M., Sundaram, R., Gupta, J. K., Kleijnen, J., & Khan, K. S. (2004). The accuracy of risk scores in predicting preterm birth--a systematic review. *Journal of Obstetrics and Gynaecology, 24*(4), 343-359. doi: 10.1080/01443610410001685439
- Hux, V. J., Catov, J. M., & Roberts, J. M. (2014). Allostatic load in women with a history of low birth weight infants: the national health and nutrition examination survey. *Journal of Women's Health (Larchmt), 23*(12), 1039-1045. doi: 10.1089/jwh.2013.4572
- King, N. M., Chambers, J., O'Donnell, K., Jayaweera, S. R., Williamson, C., & Glover, V. A. (2010). Anxiety, depression and saliva cortisol in women with a medical disorder during

- pregnancy. *Archives of Women's Mental Health*, 13(4), 339-345. doi: 10.1007/s00737-009-0139-5
- Martin, J. A., Hamilton, B. E., & Osterman, M. J. K. (2017). Births in the United States, 2016. *NCHS Data Brief*, (287), 1-8.
- Mazaki-Tovi, S., Romero, R., Kusanovic, J. P., Erez, O., Pineles, B. L., Gotsch, F., . . . Hassan, S. S. (2007). Recurrent preterm birth. *Seminars in Perinatology*, 31(3), 142-158. doi: 10.1053/j.semperi.2007.04.001
- Centers for Disease Control and Prevention, & National Center for Health Statistics. National Health and Nutrition Examination and Laboratory Protocols [1999–2006]. Retrieved December 2, 2015, from http://www.cdc.gov/nchs/nhanes/nhanes_questionnaires.htm
- Seeman, T., Gleib, D., Goldman, N., Weinstein, M., Singer, B., & Lin, Y. H. (2004). Social relationships and allostatic load in Taiwanese elderly and near elderly. *Social Science & Medicine*, 59(11), 2245-2257. doi: 10.1016/j.socscimed.2004.03.027
- Shaw, J. G., Asch, S. M., Kimerling, R., Frayne, S. M., Shaw, K. A., & Phibbs, C. S. (2014). Posttraumatic stress disorder and risk of spontaneous preterm birth. *Obstetrics & Gynecology*, 124(6), 1111-1119. doi: 10.1097/AOG.0000000000000542
- Shea, A. K., Streiner, D. L., Fleming, A., Kamath, M. V., Broad, K., & Steiner, M. (2007). The effect of depression, anxiety and early life trauma on the cortisol awakening response during pregnancy: Preliminary results. *Psychoneuroendocrinology*, 32(8-10), 1013-1020. doi: 10.1016/j.psyneuen.2007.07.006
- Soma-Pillay, P., Nelson-Piercy, C., Tolppanen, H., & Mebazaa, A. (2016). Physiological changes in pregnancy. *Cardiovasc J Afr*, 27(2), 89-94. doi: 10.5830/CVJA-2016-021
- Team, R. C. (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. .
- Voegtline, K. M., Costigan, K. A., Kivlighan, K. T., Laudenslager, M. L., Henderson, J. L., & DiPietro, J. A. (2013). Concurrent levels of maternal salivary cortisol are unrelated to self-reported psychological measures in low-risk pregnant women. *Archives of Women's Mental Health*, 16(2), 101-108. doi: 10.1007/s00737-012-0321-z
- von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gotsche, P. C., Vandenbroucke, J. P., & Initiative, S. (2008). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Journal of Clinical Epidemiology*, 61(4), 344-349. doi: 10.1016/j.jclinepi.2007.11.008
- von Versen-Hoeynck, F. M., Hubel, C. A., Gallaher, M. J., Gammill, H. S., & Powers, R. W. (2009). Plasma levels of inflammatory markers neopterin, sialic acid, and C-reactive protein in pregnancy and preeclampsia. *American Journal of Hypertension*, 22(6), 687-692. doi: 10.1038/ajh.2009.54
- Wallace, M., Harville, E., Theall, K., Webber, L., Chen, W., & Berenson, G. (2013). Neighborhood poverty, allostatic load, and birth outcomes in African American and white women: findings from the Bogalusa Heart Study. *Health & Place*, 24, 260-266. doi: <http://dx.doi.org/10.1016/j.healthplace.2013.10.002>
- Watts, D. H., Krohn, M. A., Wener, M. H., & Eschenbach, D. A. (1991). C-reactive protein in normal pregnancy. *Obstetrics & Gynecology*, 77(2), 176-180.
- Xu, J., Murphy, S. L., Kochanek, K. D., & Arias, E. (2016). Mortality in the United States, 2015. *NCHS Data Brief*, (267), 1-8.

Zipf, G., Chiappa, M., Porter, K. S., Ostchega, Y., Lewis, B. G., & Dostal, J. (2013). National health and nutrition examination survey: plan and operations, 1999-2010. *Vital and Health Statistics Series, 1*(56), 1-37.

Table 4-1: Description of measures

Variable	Source
Sociodemographic	Sociodemographic information was obtained from the household interview.
Age	Age was measured as age in months at the time of examination. Age was converted into years and was also categorized into 18 years old or less, 19-34, and 35 years old or more for analyses.
Race/ethnicity	Race was based on NHANES categories: Mexican American, other Hispanic, non-Hispanic white, non-Hispanic black, and other race including multi-racial. Race was dichotomized into non-Hispanic black and other races in this study.
Education	Education level was categorized into less than high school, high school diploma including GED, and more than high school. “Refused” or “I don’t know” was considered missing data.
Marital status	Marital status was coded as married/living with partner and widowed/divorced/separated/never married. “Refused” or “I don’t know” was considered missing data.
Poverty income ratio	This is an index for the ratio of family income to poverty threshold. The Department of Health and Human Services’ (HHS) poverty guidelines was used as the poverty measure to calculate this index. The variable was computed by dividing family income by the poverty guidelines, specific to family size, as well as the appropriate year and state. Values at or above 5.00 were collapsed to 5.00 because of disclosure concerns. The value of the variable ranges between 0-5.00.
Allostatic load	The 10 indicators included in the study were C-reactive protein (CRP) from the immune system, systolic blood pressure (SBP), diastolic blood pressure (DBP), and pulse from the cardiovascular system, body mass index (BMI), total cholesterol (TC), high-density lipoprotein (HDL), glycohemoglobin, glucose, and triglycerides from the metabolic system. Standard examination and laboratory procedures were described in the NHANES Examination and Laboratory Protocols (CDC & NCHS, 2015).
Outcome	
History of adverse birth outcomes	The study used two questions—“Did any of your children weigh less than 5 1/2 pounds (2,500 g) at birth?” and “How many of these babies were born preterm? A preterm delivery is one that occurs at 36 weeks or earlier in pregnancy.” from the Reproductive Health Questionnaire to classify participants into 2 groups. Women reporting history of delivering a premature low birth weight infant were considered as the group with history of adverse birth outcomes and those with no history of either low birth weight or preterm birth were categorized as controls. Women who answered “I don’t know” were considered as missing data.
Gestational month	We used the question –“What month of pregnancy are you in?” from the Reproductive Health Questionnaire to measure gestational month.

Table 4-2: The descriptive statistics of demographics, gestational month, and history of adverse birth outcomes

	Aim A (N=1056)		Aim B (N=665)	
	N	M±SD/%	N	M±SD/%
Age	1056	27.0±5.6	665	28.2±5.4
≤18	56	5.3	12	1.8
19-34	879	83.2	554	83.3
≥35	121	11.5	99	14.9
Poverty income ratio	988	2.4±1.7	626	2.1±1.6
Race	1056		665	
Mexican American	314	29.7	218	32.8
Other Hispanic	55	5.2	38	5.7
Non-Hispanic White	482	45.6	287	43.2
Non-Hispanic Black	145	13.7	91	13.7
Other Race - Including Multi-Racial	60	5.7	31	4.7
Education level	1055		664	
Less than high school	332	31.5	226	34.0
High school diploma including GED	216	20.5	137	20.6
More than high school	507	48.1	301	45.3
Marital status	1022			
Married/living with partner	783	76.8	511	80.0
Widowed/divorced/separated/never married	237	23.2	128	20.0
Gestational month	1056	5.6±2.2	665	5.6±2.2
Trimester	1056		665	
1st Trimester (Months 1-3)	209	19.8	134	20.2
2nd Trimester (Months 4-6)	439	41.9	270	40.6
3rd Trimester (Months 7-9)	408	38.7	261	39.2
History of adverse birth outcomes			665	
Yes			31	4.7
No			634	95.3

Table 4-3: The summary of the 10 indicators in terms of their gestational patterns

Indicators	Patterns observed	Statistically significant changes across gestation by month or trimester?	Consistent with known pregnancy physiology?
Pulse	Near linear. Pulse increased with the progression of pregnancy but had a slightly decrease at late gestation.	Yes. Gestational month was significantly correlated with pulse ($\beta=0.07$, $P<0.001$).	Yes. The early increase in ventricular wall muscle mass and end-diastolic volume contribute to an increase in stroke volume and heart rate in pregnancy (Soma-Pillay, Nelson-Piercy, Tolppanen, & Mebazaa, 2016).
SBP	Inverted bell curve. There was a decrease in SBP from early gestation to mid-gestation and then SBP increased to the late gestation.	Yes. SBP at the first and third trimester were significantly higher than the second trimester ($\beta=0.12$, $P=0.033$; $\beta=0.17$, $P<0.001$).	Yes. The smooth muscle relaxation and overall vasodilation caused by elevated progesterone contribute to the decreased SBP in the first and second trimester. The highly increased plasma volume in the third trimester causes an increase in SBP (Soma-Pillay et al., 2016).
DBP	Inverted bell curve. DBP decreased up to the middle of gestation and then increased up to the late gestation.	Yes. DBP at the first and third trimester were significantly higher than the second trimester ($\beta=0.32$, $P<0.001$; $\beta=0.09$, $P<0.001$).	Yes. The explanation for the pattern of DBP is the same as SBP (Soma-Pillay et al., 2016).
BMI	Inverted bell curve. BMI decreased in the first month, remained steady in the following 4 month, and then increased up to the late gestation.	Yes. BMI at the third trimester were significantly higher than the first and second trimester ($\beta=0.25$, $P<0.001$; $\beta=0.12$, $P<0.001$).	Yes. Pregnant women lose their weights at early gestation maybe due to nausea and vomiting. After the symptoms reduce in the second trimester, their body weights increase.
TC	Near linear. There was a steady increase in TC as pregnancy progresses.	Yes. Gestational month was significantly correlated with TC ($\beta=0.15$, $P<0.001$).	Yes. The increase in TC levels is mainly due to increased synthesis by the liver and decreased lipoprotein lipase activity, resulting in decreased catabolism of adipose tissue. Changes in lipid metabolism accommodate the needs of the developing fetus. Increased TC levels provide for the mother's energy needs while glucose is spared for the fetus (Soma-Pillay et al., 2016).
HDL	Bell curve. HDL increased in the first half of	Yes. HDL levels at the first and third trimester were	Yes. The explanation for the pattern of HDL is the same as TC (Soma-Pillay et al., 2016).

	pregnancy and then fall in the third trimester, but the level was still higher than early pregnancy.	significantly lower than the second trimester ($\beta=-0.37$, $P<0.001$; $\beta=-0.05$, $P=0.042$).	
CRP	Bell curve. As pregnancy progressed, CRP levels fluctuated with the peak levels reached at gestational month 4 and 7.	No. CRP levels at the first and third trimester were not significantly lower than the second trimester ($\beta=-0.03$, $P=0.620$; $\beta=-0.02$, $P=0.541$).	Partially consistent. CRP is known to be slightly elevated during pregnancy, due to the maternal inflammatory reaction to the pregnancy, but there is no consistent change in CRP levels with gestational age (von Versen-Hoeynck, Hubel, Gallaher, Gammill, & Powers, 2009; Watts, Krohn, Wener, & Eschenbach, 1991).
Glycohemoglobin	Inverted bell curve. Glycohemoglobin fell in the half of pregnancy and then rose to the late pregnancy.	Yes. Glycohemoglobin levels at the first and third trimester were significantly higher than the second trimester ($\beta=0.33$, $P<0.001$; $\beta=0.14$, $P<0.001$).	Yes. The pattern is the result of increased insulin secretion and increased insulin sensitivity in early pregnancy, followed by progressive insulin resistance in the third trimester (Soma-Pillay et al., 2016).
Glucose	Inverted bell curve. Glucose decreased in the middle of gestation and then increased to the late gestation.	Yes. Glucose levels at the first and third trimester were significantly higher than the second trimester ($\beta=0.24$, $P<0.001$; $\beta=0.05$, $P=0.045$).	Yes. The explanation for the pattern of glucose is the same as glycohemoglobin (Soma-Pillay et al., 2016).
Triglycerides	Near linear. There was a steady increase in triglycerides with the progression of pregnancy.	Yes. Gestational month was significantly correlated with triglycerides levels ($\beta=0.17$, $P<0.001$).	Yes. The explanation for the pattern of triglycerides is the same as TC (Soma-Pillay et al., 2016).

Note. SBP, systolic blood pressure; DBP, diastolic blood pressure; BMI, body mass index; TC, total cholesterol; HDL, high-density lipoprotein; CRP, C-reactive protein

Table 4-4: The associations between history of adverse birth outcomes and the allostatic load index

	Model 1 (Unadjusted)				Model 2 (Adjusted)			
	OR	95% CI	<i>P</i>	Pseudo <i>R</i> ²	OR	95% CI	<i>P</i>	Pseudo <i>R</i> ²
ALI	1.24	1.0002-1.52	0.044	0.10	1.25	1.002-1.56	0.044	0.18
Aged 18 or less ^a					1.27	0.065- 7.83	0.832	
Aged 35 or more ^a					0.73	0.17- 2.26	0.624	
Non-Hispanic Black ^b					1.84	0.67- 4.52	0.204	
Poverty income ratio					0.76	0.55- 1.02	0.085	

Note. ALI, allostatic load index

^a Aged 19-34 years was the reference; ^b Other races was the reference.

Pseudo *R*² is Nagelkerke's R-squared.

Table 4-5: The cut-off points, sensitivities, and specificities of the allostatic load index with history of adverse birth outcomes as the outcome

Cut-off point as number of high risk quartiles (0-9)	Sensitivity	Specificity
0	1	0.12
1	0.79	0.33
2	0.62	0.59
3	0.31	0.79
4	0.10	0.90
5	0.10	0.96
6	0.03	0.986
7	0	0.994
8	0	0.998
9	0	1

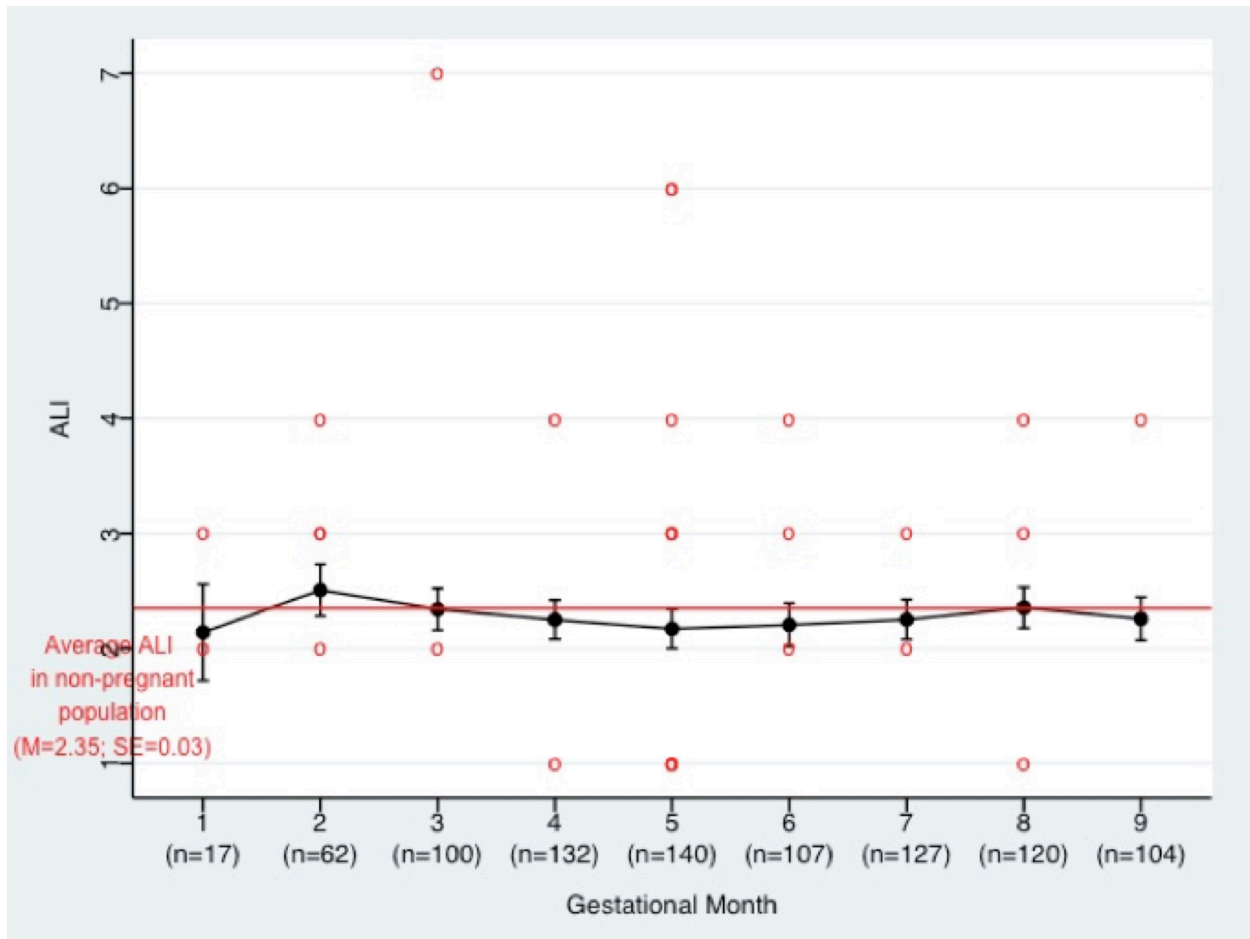
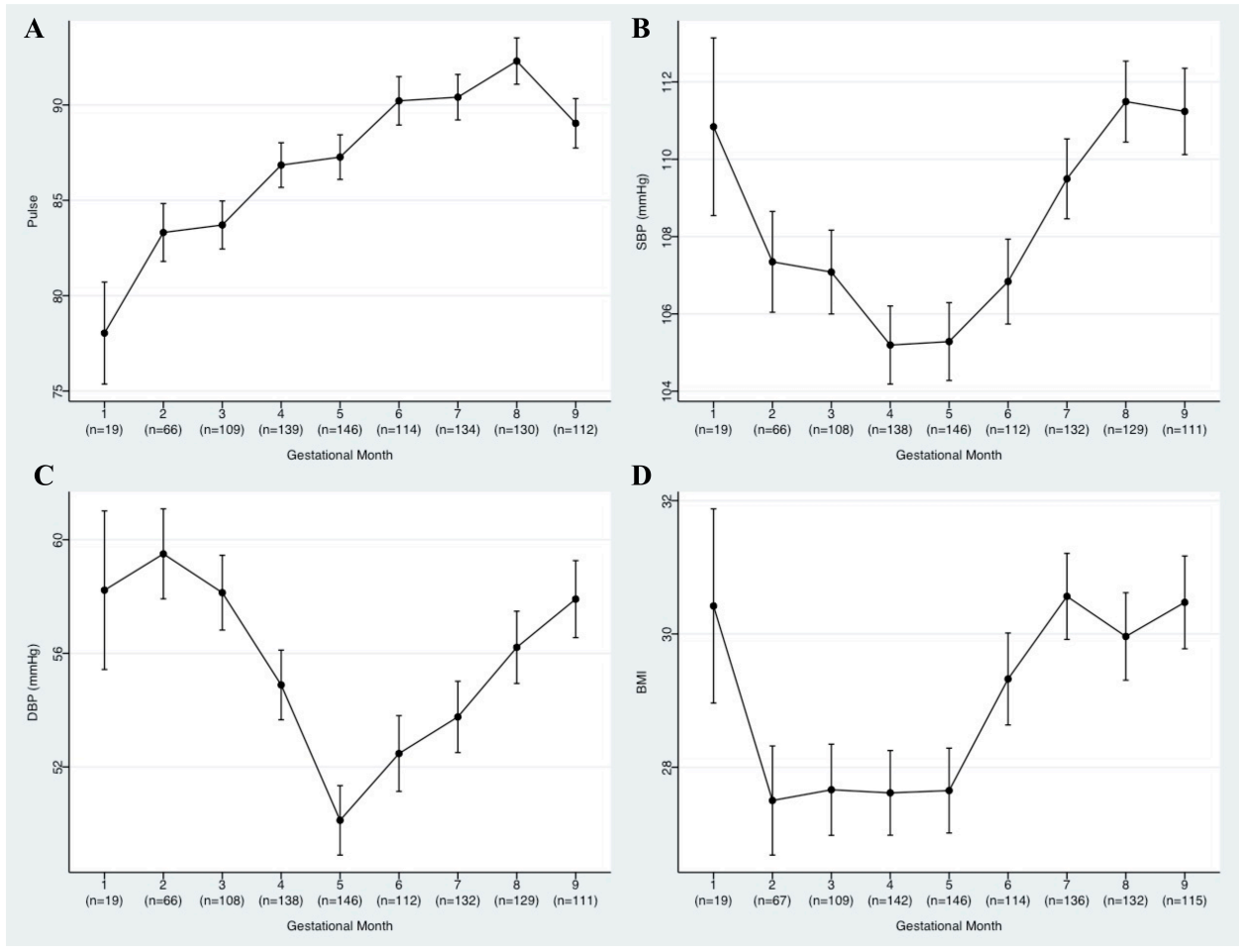


Figure 4-1: The adjusted means and standard errors of the allostatic load index by the count-based method across gestational month with allostatic load index scores in pregnant women with prior adverse birth outcomes plotted and the average allostatic load index in the non-pregnant population from the NHANES 2001-2006 data shown for comparison
Note. Age, race, and poverty income ratio were adjusted for.



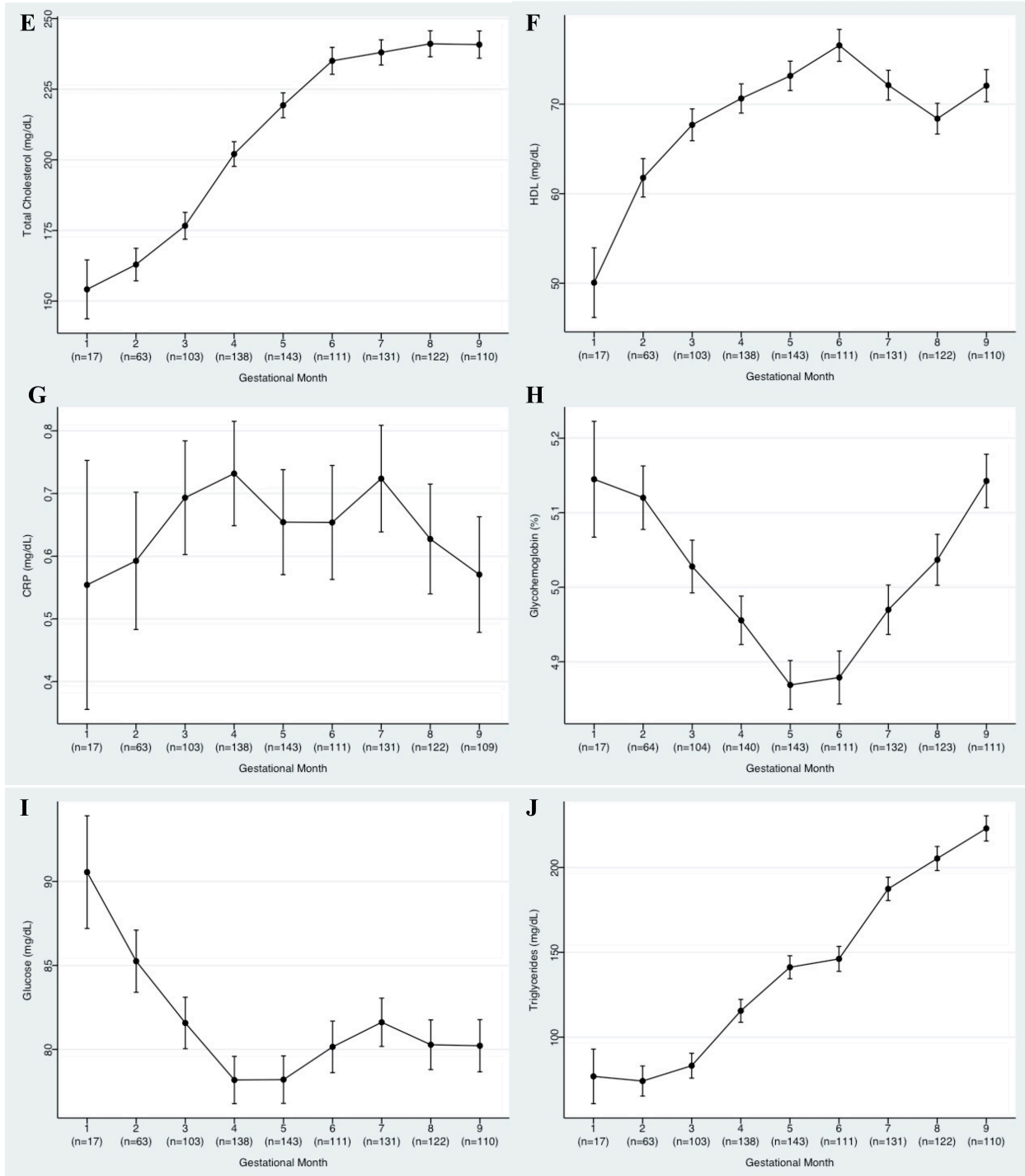


Figure 4-2: The adjusted means and standard errors of the 10 allostatic load indicators across gestational month

Note. Age, race, and poverty income ratio were adjusted for.

Chapter 5: Conclusion

The summary of findings

This dissertation project is comprised of 3 parts that addressed 3 aims. In the first part of the dissertation, we applied the allostatic load (AL) theory to perinatal outcomes research and proposed a theoretical model to understand AL as a potential biological mechanism for the impact of maternal chronic stress such as posttraumatic stress disorder (PTSD), depression, and child maltreatment on adverse perinatal outcomes. The second part of the dissertation was a secondary analysis of data from the National Health and Nutrition Examination Survey (NHANES). We focused on women of reproductive age and created AL indices (ALI) using the 5 scoring methods including the count-based, Z-Score, multivariable logistic regression, factor analysis, and grade of membership (GOM) method. We then examined the predictive performances of each ALI score for predicting general health status, diabetes, and hypertension and made recommendations for which scoring method is optimal for use. The third part of the dissertation was also a secondary analysis of data from the NHANES, which focused on pregnant women. We plotted gestational curves to describe changes in each AL indicator and the ALI score by the optimal scoring method, with the average ALI score in the non-pregnant population as a comparison. We also tested the association between the ALI score and history of delivering a premature low birth weight infant as a proxy for a subsequent adverse birth outcome on this pregnancy. We then evaluated the predictive performance of the ALI score for predicting prior adverse birth outcomes.

The theoretical model construction

The proposed perinatal AL model can be depicted as follows: Child maltreatment, intimate partner violence (IPV), and low socioeconomic status (SES) as maternal chronic stress increase the risk for developing maternal PTSD. In response to maternal PTSD, primary mediators (e.g., cortisol, epinephrine, norepinephrine) from the neuroendocrine system are produced and cause a cascade of effects on secondary outcomes that are dysregulations in immune, metabolic, and cardiovascular systems. Those cumulative physiological dysregulations across multiple systems either directly lead to quaternary outcomes (i.e., preterm birth and low birth weight) or indirectly increase the risk of tertiary outcomes (i.e., preeclampsia, gestational diabetes, infection, premature rupture of membrane) that eventually result in quaternary outcomes. Comorbid depression, pregnancy-related distress, and risk behaviors interact with maternal PTSD to enhance the effect of maternal PTSD on adverse birth outcomes. The theory places AL as a mediator and places comorbid depression, pregnancy-related distress, and risk behaviors as moderators.

The AL scoring issue

The ALI score by the logistic regression method had the best predictive performances with regard to general health status, diabetes, and hypertension compared to the other 4 scoring approaches. The differences among the 5 summary scores were not major, suggesting that the advantages of one method over another are relatively subtle. We also found the logistic regression method performed better when some AL components have much stronger associations with specific health outcomes than the other AL components.

Given that the logistic regression method had the strongest predictive validity compared to the other methods, the logistic regression method is optimal for use. But this method relies on

subsequent outcome information. It is challenging to compare ALI scores across different outcomes. It may also not be applicable when the outcome information is unknown. Under the above conditions, the count-based method may be a good alternative. The predictive performance of the ALI score by this method is similar to the Z-Score, factor analysis, and GOM method. Compared to the other scoring approaches, the count-based method has its own strengths. It is the most commonly used method in previous AL studies. Given that the ALI score by this method is the number of indicators of risk for adverse health outcomes, it is easy to compute and interpret. Therefore, if the outcome information is known or consistent across different contexts, we recommended the logistic regression method; otherwise the count-based method may be a good alternative in terms of predictive validity, feasibility, and interpretability.

The pregnancy AL

The gestational patterns of AL indicators were consistent with known pregnancy physiology. Except for C-reactive protein, significant differences were detected for each indicator across different gestational month, suggesting that gestational age needs to be taken into consideration when scoring AL during pregnancy. The ALI score at each gestational month was not different from the average score in the non-pregnant population. It suggested that measuring AL at any gestational time point would reflect women's true physiological functions, but gestational age should be taken into account when scoring AL.

We found poor predictive performance of the ALI score for predicting prior adverse birth outcomes. The ALI score failed to identify the majority of women reporting history of adverse birth outcomes, which is similar to the performances of existing risk scores and technical screening tools. The low accuracy may be due to the multifactorial nature of adverse birth outcomes. Although there was a statistically significant association of the ALI score with history

of adverse birth outcomes, the effect size was small. Thus, the AL summary measure is not sufficiently sensitive to use as a single predictor for the risk of adverse birth outcomes.

Implications for research

This dissertation project has a few implications for research. First, the synthesized theoretical model may be useful for its potential to advance perinatal research. It may provide a theoretical groundwork for future research to examine AL as a potential pathway for the link between maternal chronic stress and adverse perinatal outcomes. Based on the theoretical model, prospective studies with multiple physiological indicators, measures of maternal chronic stress (i.e., PTSD, child maltreatment, IPV, lower SES, depression, pregnancy-related distress, risk behaviors), and following up with women's perinatal outcomes may add new knowledge about the complex mechanisms leading to adverse perinatal outcomes. The improved understanding of the complex multifactorial causal pathways for adverse perinatal outcomes assists in developing highly accurate screening assessments for adverse perinatal outcomes.

Second, the second part of the dissertation advanced AL research by validating different scoring approaches with a nationally representative sample and recommending the optimal scoring method for use. Future research could use the recommended scoring method to construct an ALI score. A commonly accepted scoring approach could facilitate comparisons of findings across different AL studies.

Third, the third part of the dissertation contributed to perinatal research in terms of measuring AL during pregnancy. Future research should take gestational age into account when scoring AL in pregnant women and test whether pregnancy AL could mediate the impact of maternal chronic stress on adverse perinatal outcomes.

These 3 parts of the dissertation together may lay critical theoretical and methodological groundwork for future research on perinatal outcomes. Study #1 proposed a theoretical model which may be a theoretical basis for examining the link between maternal chronic stress and adverse perinatal outcomes. Study #2 resulted in recommendations of the optimal AL scoring method for use. Study #3 added new knowledge about scoring AL in pregnancy as well as the validity of pregnancy AL for predicting adverse birth outcomes. Taken together, results of these 3 dissertation studies would be underpinnings for a prospective perinatal outcome study.

Implications for clinical practice

The dissertation project is a preliminary study and the proposed theoretical model was not tested. More work is needed to validate the model and test if AL can be used as a sensitive screening assessment for adverse perinatal outcomes. But this project may still have a few potential implications for clinical practice. First, after the proposed theoretical model is well validated, it may be applied to clinical practice to help health care providers perceive the association between maternal stress, health alterations that accumulate and interact, and adverse birth outcomes. This may lead to actions to decrease stress reactions from health care itself and from having low resources.

Second, this project may advance early identification of women at risk for adverse perinatal outcomes. Brief, effective screening tools for maternal chronic stress such as PTSD, depression, pregnancy-related distress, history of child maltreatment, and IPV could be designed and included in prenatal care to identify vulnerable women. The study of these physiological indicators may provide us not only with insights into the biological mechanisms that contribute to adverse perinatal outcomes but also with screening tools to identify women at risk for adverse perinatal outcomes. Given that this study found high specificity and negative predictive value,

the summary measure of the accumulated physiological dysregulations may be used to identify pregnant women who are likely to be at low risk for adverse perinatal outcomes. This would avoid unnecessary prenatal care and medical costs. Evaluating women's levels of physiological indicators may also be added to existing screening assessments to improve the accuracy for identifying women subsequently having an adverse perinatal outcome. Most of AL indicators such as BMI, blood pressure, and pulse are routine clinical assessments and thus are feasible to be measured. The logistic regression method can be used through computer software and the count-based ALI as an alternative measure can be easily calculated by hand. After the majority of women who subsequently report adverse perinatal outcomes can be identified, they can receive appropriate care to mitigate their pregnancy risks. An acceptably accurate screening tool will allow health care providers to objectively assess a pregnant woman's overall risk and provide health care according to her risk level.

Third, this project may provide valuable insights into preventions and interventions that aim to reduce adverse perinatal outcomes. Considering the effects of maternal chronic stress on mother and infant health, health care providers including obstetricians and midwives, need to be aware of those chronic stressors, especially PTSD and depression. Appropriate mental health care and education about how to identify and reduce stress should be provided for high-risk women in the prenatal period in order to address stress dysregulation and improve both maternal well being and infant outcomes. Since mental health care services may be nonexistent or limited and predominantly hospital based (Saxena et al., 2011), consideration should be given to developing or strengthening local mental health care services that will serve women during the perinatal period. Additionally, women at risk for adverse perinatal outcomes that are identified by highly accurate screening tools should be targeted for more intensive supervision and

appropriate levels of prenatal care. Effective interventions such as some stress reduction techniques should also be designed and evaluated to prevent those adverse perinatal outcomes. If an effective intervention to prevent or delay adverse perinatal outcomes were lacking, any screening test would likely be of no clinical utility. Only when the use of a highly accurate screening tool and subsequent effective interventions has been evaluated to significantly reduce the occurrence of adverse perinatal outcomes should the screening tool be included in routine prenatal care.

Limitations and strengths

Limitations

This project had some limitations. First, the proposed theoretical model was not tested because of the data limitation. Second, we validated different AL scoring approaches in women of reproductive age. Our findings may not be generalized to the male or elder population. But focusing on a more homogeneous female population may enhance the reliability of our findings. Third, data on physiological indicators from the primary mediating neuroendocrine system are lacking in the NHANES database. The ALI was constructed without indicators from the neuroendocrine system, relying solely on the indicators of secondary dysregulations for the scoring, which may reduce the predictive performances of the summary measure on health outcomes. Fourth, due to the cross-sectional study design of the NHANES, data on the outcome variables and AL indicators were collected at the same time. This may affect the predictive performances of ALI scores for predicting general health status, diabetes, and hypertension. Fifth, because of the cross-sectional study design, no information on subsequent perinatal outcomes of the current pregnancy is available. Only history of delivering premature low birth weight infants was measured, which was considered as a proxy adverse birth outcome to be

included as the outcome variable for analysis. This may decrease the predictive validity of pregnancy ALI scores for predicting adverse birth outcomes.

Strengths

The study had strengths as well. First, the NHANES database includes numerous physiological measures from multiple body systems, which provides adequate AL indicators for our study to construct the ALI. Second, data on pregnant women is available in the NHANES database. The multi-system physiological measures during pregnancy and a large sample size enabled us to construct a pregnancy ALI. Third, data was collected with standardized procedures and protocols to assure that the data for this analysis is of high quality in terms of reliability and validity.

Future research directions

To apply the AL theory to perinatal research, there are several research directions for researchers to focus on. First, despite constructing a theoretical model for perinatal outcomes research, prospective studies with multiple biological and psychosocial measures of stress and following up women's perinatal outcomes are needed to validate and confirm the theoretical model.

Second, due to the data limitations, the second and third parts of the dissertation are only preliminary studies that aimed to address some methodological issues (the AL scoring issue and the validation of AL in pregnancy). Future research is still needed to address those methodological issues before testing the theoretical model. Prospective longitudinal studies using a full complement of physiological indicators to operationalize the AL and using different scoring approaches in different age- and gender-specific populations will be required to validate

the recommendations made based on this secondary analysis. Age- and gender-specific population norms for the ALI score by the optimal scoring method will be also needed.

Third, although the count-based measure may a good alternative for use, it has the limitation of making the ALI sample-specific. To address the limitation, we should create clinical risk cut-off points based on national standards instead of risk quartiles of the sample distribution to count the total number of indicators of risk for poor health. Thus, further work on establishing population norms in terms of age, sex, race, etc. is needed.

Fourth, future research with adequate sample of pregnant women is needed to measure physiological indicators at different gestational time points and follow up their perinatal outcomes to examine the validity of AL in pregnancy. Accordingly, a cut-off point of pregnancy ALI score can be recommended after research on data where the outcome is adverse birth outcomes subsequent to the current pregnancy instead of prior adverse birth outcomes. Population-based cut-off points can be also determined for each AL indicator at each gestational month to establish population norms for pregnancy.

Fifth, given the low accuracies of the ALI score, existing risk scoring systems, and technical assessments, combining them together may improve the overall accuracy compared with single approach alone. Future research needs to explore whether a comprehensive screening approach would have a high predictive performance for adverse perinatal outcomes.

Lastly, what next can be feasibly done by us is that we could use the NHANES data to identify cut-off points for each AL indicator at each gestational month. Given that the NHANES has a nationally representative sample of pregnant women with physiological data collected with standardized procedures and protocols, population norms for pregnancy can be established for use in research and clinical practice.