

Algorithms for Weight Measures using VHA Administrative Data: A Simpler Approach may be Sufficient

Wyndy L. Wiitala, PhD¹, Richard Evans, MS¹, Ann Annis, RN, PhD^{1,2}, Jennifer A. Burns, MHSA¹, Michelle B. Freitag MPH¹, Susan D. Raffa PhD^{3,4}, Michael G. Goldstein³, Laura J. Damschroder, MS, MPH¹

¹ Center for Clinical Management Research, VA Ann Arbor Healthcare System, Ann Arbor, MI
² College of Nursing, Michigan State University, East Lansing, MI
³ National Center for Health Promotion and Disease Prevention, Veterans Health Administration, Durham, NC
⁴ Department of Psychiatry & Behavioral Sciences, Duke University School of Medicine, Durham, NC

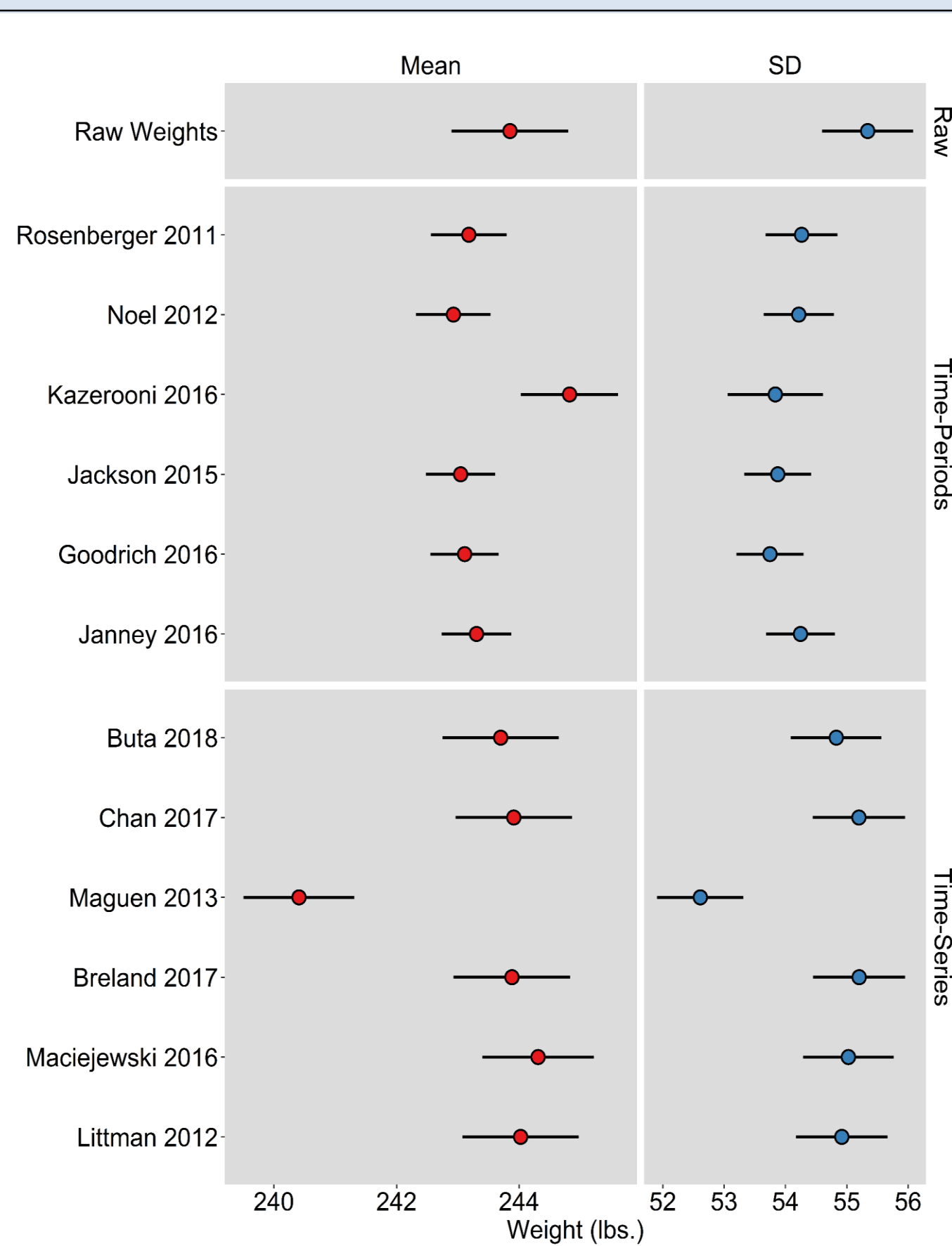


Statistical Properties

Descriptive Statistics for 12 algorithms

Algorithm	N Pts. Retained (% of Raw)	N Weights Retained (% of Raw)	Range (Min, Max)
Raw Weights	99,012 (100)	2,718,522 (100)	(0, 1495)
Time-Periods			
Rosenberger 2011	79,468 (80)	295,702 (11)	(0, 1397)
Noel 2012	99,012 (100)	966,547 (36)	(70, 651)
Kazerooni 2016	53,738 (54)	161,214 (6)	(1, 654)
Jackson 2015	96,248 (97)	281,687 (10)	(75, 614)
Goodrich 2016	97,461 (98)	244,878 (9)	(82, 500)
Janney 2016	97,487 (98)	245,073 (9)	(72, 614)
Time-Series			
Buta 2018	95,517 (96)	2,655,380 (98)	(62, 290)
Chan 2017	98,558 (99)	2,706,887 (99)	(52, 743)
Maguen 2013	98,756 (99)	2,270,847 (84)	(72, 608)
Breland 2017	99,012 (100)	2,717,137 (99)	(75, 651)
Maciejewski 2016	99,011 (99)	2,586,266 (98)	(52, 621)
Littman 2012	98,558 (99)	2,683,266 (98)	(75, 598)

95% CI of Means and Standard Deviations



Bootstrapped 95% CI for Mean, SD by Algorithm. Data compiled from 100 iterations of 1,000 randomly sampled patients with replacement.

Weight Processing by Algorithm and Type of Algorithm. Time-Period algorithms define windows around time points; Time-Series use all available weights. Algorithms using specific time periods remove > 50% of the data. Implausible measurements remain despite algorithm implementation.

Summary

- Systematic review of 39 studies from 2008-2018 that used VHA EHR weight data as an outcome.
- Reconstructed 29 algorithms and applied 12 using data from patients enrolled in MOVE! in 2016.
- Resulting weight measures were compared at the patient (N=99,012) and facility (N=129) levels.
- Applying basic cut-offs that require fewer computing resources and are cognitively simpler may be sufficient for many studies (e.g., examining point estimates, one-year weight change).
- Other analyses (e.g., trajectories, facility-level comparisons) may require more nuanced approaches.

Weight Change Outcomes

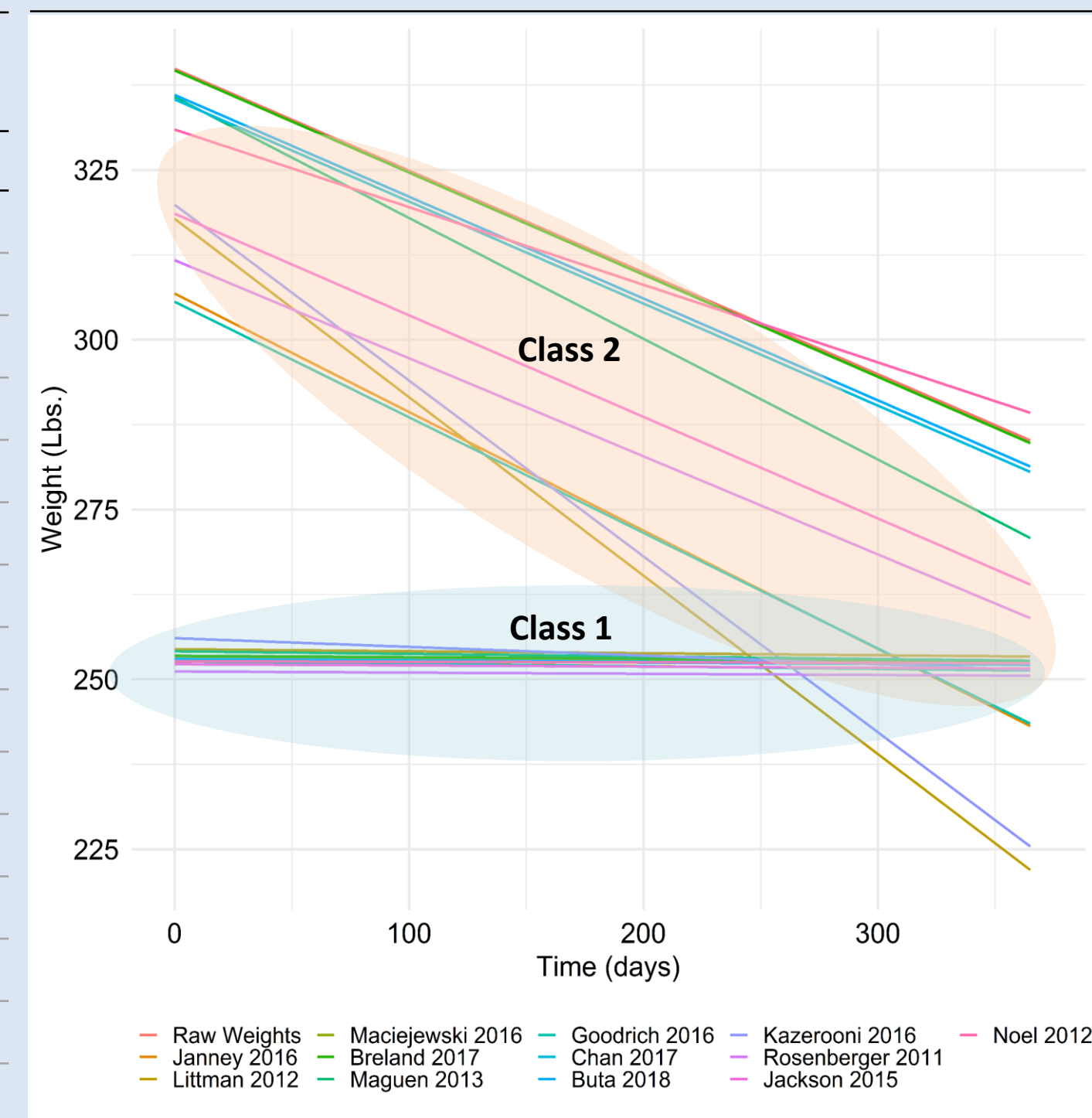
Common Weight Change Outcome Metrics

Algorithm	N Pts. Retained	Weight Loss ≥5%	Avg Weight Change (lbs.)	
	N (%)	N (%)	Mean (SD)	(Min, Max)
Raw Weights	73,233 (74)	13,814 (19)	-2.90 (17)	(-334, 274)
Time-Periods				
Rosenberger 2011	49,177 (50)	9,740 (20)	-3.10 (18)	(-1176, 270)
Noel 2012	69,508 (70)	12,376 (18)	-2.45 (16)	(-321, 164)
Kazerooni 2016	53,738 (54)	10,257 (19)	-3.64 (15)	(-343, 369)
Jackson 2015	51,619 (52)	10,081 (20)	-3.39 (16)	(-262, 137)
Goodrich 2016	69,882 (71)	13,207 (19)	-2.90 (16)	(-223, 271)
Janney 2016	69,966 (71)	13,240 (19)	-2.93 (17)	(-321, 271)
Time-Series				
Buta 2018	71,244 (72)	13,420 (19)	-2.87 (16)	(-223, 184)
Chan 2017	73,010 (74)	13,604 (19)	-2.80 (16)	(-321, 207)
Maguen 2013	62,075 (63)	8,050 (13)	-1.59 (11)	(-134, 81)
Breland 2017	73,187 (74)	13,771 (19)	-2.85 (16)	(-265, 141)
Maciejewski 2016	69,429 (70)	13,597 (20)	-2.94 (16)	(-265, 141)
Littman 2012	72,294 (73)	13,132 (18)	-2.35 (15)	(-164, 145)

One-Year Weight Loss Metrics by Algorithm. Metrics are generally similar across algorithms, but implausible values of weight change remain.

Weight Trajectories

Latent Class Mixed Model



Group-Based Trajectory Modelling by Algorithm. Predictions from latent class linear mixed models identify 2 trajectories; algorithm choice impacts predictions in Class 2.

Site-Level

- We examined and ranked 129 facilities by the percent of patients with at least 5% weight loss by algorithm.
- Facility percentages varied by algorithm (minimum: 3% - 13%; maximum: 22% - 28%).
- The median difference for facilities across methods was 8%; differences ranged from 3 - 17%.
- The median rank difference within facility across methods was 46; differences ranged from 5 to 111.
- Time-series algorithms may result in more consistent rates; results are mixed for time-period algorithms.