

Applying RE-AIM to Evaluate the External Validity of Weight Gain Prevention Interventions in Young Adults: A Systematic Review

Debra Haire-Joshu, PhD; Alexandra B. Morshed, PhD, MS; Allison Phad, MPH; Shelly Johnston, MPH, RD; Rachel G. Tabak, PhD, RD

ABSTRACT

Objectives: The purpose of this review was to use RE-AIM (Reach, Effectiveness, Adoption, Implementation, Maintenance) to assess the extent to which weight gain prevention studies targeting young adults reported on elements of external validity.

Design: Systematic review.

Eligibility Criteria: Articles of interest included a lifestyle/behavioral intervention targeting weight gain prevention. Eligibility criteria included the following: study design of randomized controlled trials, quasi-randomized control trials, or natural experiments; average participant age between 18 and 35 years; study duration of at least 12 months; and published in English between January 2008 and May 2018. Studies had to report weight or body mass index as a measured outcome and were excluded if they were paired with smoking cessation programs, were conducted in specific groups (ie, pregnant women, breast cancer survivors), or were follow-ups to weight loss studies.

Study Selection: After removing duplicates, the search yielded 11 426 articles. Titles and abstracts were screened by 1 reviewer; 144 articles were assessed in a full-text review by 2 reviewers. Discrepancies were resolved by consensus. Nine studies (13 articles) were included in the review.

Main Outcomes Measure(s): Reported elements of the RE-AIM framework.

Results: A total of 9 studies met the selection criteria. All studies lacked full reporting on external validity elements. Of the total of 60 RE-AIM reporting criteria, 8 were reported by all 9 studies, 26 criteria were reported by fewer than 4 studies, and 22 criteria were not reported by any of the studies.

Discussion: There remains inadequate reporting of elements of external validity and generalizability in weight gain prevention studies. This is a significant scientific constraint that limits the information required to disseminate and implement prevention of weight gain interventions for population impact. Standardized reporting may be needed to ensure results that demonstrate not only internal validity but also external validity and generalizability are needed to promote public health impact.

KEY WORDS: external validity, RE-AIM framework, weight gain prevention, young adults

Author Affiliations: Center for Diabetes Translation Research (Drs Haire-Joshu and Tabak and Mss Phad and Johnston), Center for Obesity Prevention and Policy Research (Dr Haire-Joshu), and Prevention Research Center (Drs Morshed and Tabak), the Brown School, Washington University in St Louis, St Louis, Missouri.

This work was made possible with support from the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) and the National Heart Lung and Blood Institute (NHLBI) of the National Institutes of Health (NIH) under award numbers P30DK092950, T32HL130357, R01DK121475-01, and R01HL143360-01. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIDDK, NHLBI, or NIH.

The authors thank Aishat Gambari (A.G.), MPH, for contributing to the evidence acquisition for this review.

All authors declare that they have no conflicts of interest.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (<http://www.JPHMP.com>).

Young adulthood is a period of high risk for excessive weight gain and development of obesity, representing an important target of intervention. On average, US adults gain 0.5 to 1.0 kg per year, resulting in an average 13-kg weight gain from early to middle adulthood.^{1,2} This average

This is an open access article distributed under the Creative Commons Attribution License 4.0 (CCBY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Correspondence: Debra Haire-Joshu, PhD, Center for Diabetes Translation Research and Center for Obesity Prevention and Policy Research, the Brown School and the School of Medicine, Washington University in St Louis, 1 Brookings Dr, Campus Box 1196, St Louis, MO 63130 (djoshu@wustl.edu).

Copyright © 2020 The Authors. Published by Wolters Kluwer Health, Inc.

DOI: 10.1097/PHH.0000000000001159

annual weight gain doubles the prevalence of obesity in the population, as young adults advance rapidly toward overweight or obesity by middle age,² and is leading to a decrease in life expectancy in the United States. Woolf and Schoemaker³ reported that, between 1999 and 2017, age-adjusted midlife mortality rates for obesity increased by 114.0% (from 1.3 deaths/100 000 to 2.7 deaths/100 000). Midlife is a particularly susceptible period for young women, due to excess weight gain and retention during pregnancy and postpartum, and further increases the disparities among racial subgroups.⁴ Risk for excessive weight gain in early adulthood is also associated with early onset of diabetes, cardiovascular, and related chronic diseases, impacting quality of life and health care costs.^{2,5}

Early weight gain prevention trials (1985-2011) were generally designed to test whether the intervention was or was not efficacious. Consistent with this focus on standard reporting elements for internal validity, reports of study findings primarily describe efficacy.^{6,7} Brought on by a lack of external validity, calls were made for more generalizable studies and the reporting on elements of external validity. External validity incorporates a better understanding of the generalizability of interventions across different populations, settings, and variations in treatment,⁸ which is needed to assess how well the research translates into practice.⁹ Reviews have recommended improvements in reporting on external validity components that influence dissemination and scale-up of interventions aimed at preventing obesity among this high-risk age group.¹⁰ Going beyond reporting, research with more relevance and generalizability is needed to impact chronic disease burden at a population level.^{11,12} To understand the gaps in external validity, it is important to review the extent to which current studies report generalizable findings. It is unclear whether weight gain prevention studies conducted since these prior reviews, and during times when calls for attention to balanced reporting were more prominent, have comprehensively addressed elements of both internal and external validity in describing their findings.

There are several approaches to guide and assess the balance of internal and external validity in study planning, execution, and reporting of study findings.¹³⁻¹⁷ The RE-AIM (Reach, Effectiveness, Adoption, Implementation, Maintenance) planning and evaluation framework¹⁸ guides the reporting of essential program elements addressing external validity that may improve the sustainable adoption and implementation of effective, generalizable, and evidence-based interventions. It has been used extensively over the past 2 decades in public health and health behavior change research to report on contextual factors related to

external validity of interventions.^{15,19} The purpose of this review was to use RE-AIM to assess the extent to which weight gain prevention studies targeting young adults reported on elements of external validity.

Methods

Search strategy and study selection

We conducted a systematic literature review of weight gain prevention studies published in peer-reviewed journals in the 10-year period from January 2008 to May 2018. We chose to focus on studies publishing results during this time period because of the heightened attention to the importance of generalizability and expanded transparency in reporting external validity during this period. Databases searched included Scopus, Web of Science, EBSCOhost, and PubMed. A complete search strategy can be found in Supplemental Digital Content Appendix Table 1 (available at <http://links.lww.com/JPHMP/A650>). In summary, search terms were broad and included combinations, truncations, and synonyms of “weight,” “weight maintenance,” “young adult,” “lifestyle,” “behavioral,” and “intervention.” The search was limited to English. Additional studies were retrieved from reference lists of relevant studies; articles related to the included studies (ie, protocol papers, formative research) were also retrieved. Studies of interest included randomized controlled trials, quasi-randomized control trials, and natural experiments. Studies needed to include a behavioral or lifestyle intervention targeting weight gain prevention, a comparison group, and weight or body mass index (BMI) as a measured outcome. Excluded studies had an average participant age greater than 35 years or included interventions targeting pregnant women, paired weight gain prevention with smoking cessation programs, were conducted in specialized groups (ie, breast cancer survivors), or were follow-ups to weight loss studies. Since this review was concerned with annual weight gain, included studies had to be at least 12 months in duration, including length of intervention and follow-up. This review has been registered at PROSPERO (International Prospective Register of Ongoing Systematic Reviews, <http://www.crd.york.ac.uk/prospero>, CRD42018091824).²⁰

After duplicates were removed, the initial search yielded 11 426 studies (Figure). Titles and abstracts were screened by 1 reviewer (S.J.); 144 full-text studies were included for full-text review and assessed for inclusion by 2 reviewers (A.P. and S.J.). Discrepancies between reviewers were resolved by consensus. From these, 9 studies (from 13 articles) were eligible and included in the review.²¹⁻³³

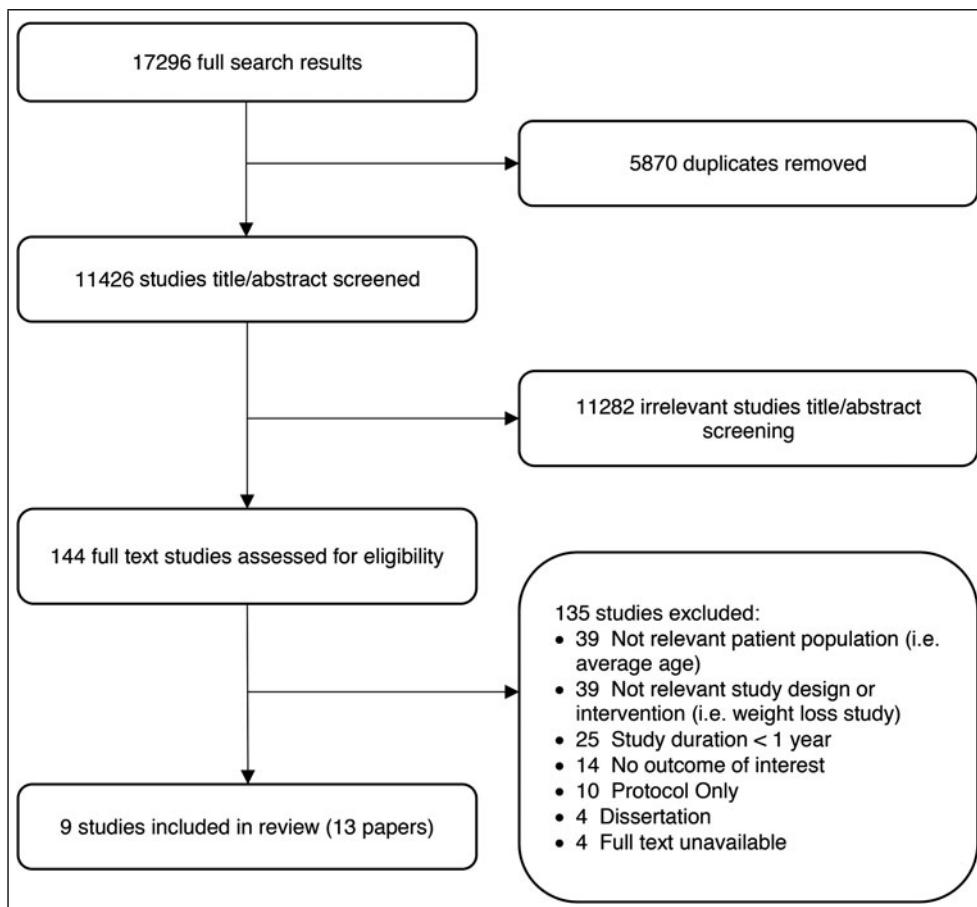


FIGURE PRISMA Flow Chart—Process of Inclusion of Studies

Data collection

Two reviewers (A.P. and A.G.) utilized the Cochrane data collection form to extract data from the 9 selected studies.^{21–29} The form was piloted on one study and then adjusted and used for those remaining. The following data were extracted: study details (authors, year, sponsorship source, country of publication), methods (design, aim of study/intervention, statistical methods, units of randomization/analysis), population characteristics (inclusion/exclusion criteria, baseline group differences, demographics), intervention details (setting, theoretical basis, content, providers, duration), and outcomes (weight, BMI, behavioral). Data were extracted for all articles by both reviewers, and discrepancies were resolved by consensus with the study team. Study screening and data extraction were completed using Covidence systematic review software.³⁴

Following extraction, 2 reviewers (R.G.T. and A.M.) used the RE-AIM framework to assess the extent of reporting in the included studies on

study elements related to internal and external validity and translation potential. The reviewers used the data extraction tool developed by Harden et al,³⁵ designed specifically for conducting systematic reviews using RE-AIM. The tool measures multiple indicators for each RE-AIM element at multiple levels (ie, individual, provider, organizational): (1) reach (eg, description of target population), (2) efficacy/effectiveness (eg, use of intent to treat), (3) adoption (eg, method to identify setting, staff participation rate), (4) implementation (eg, timing and duration of contacts), and (5) maintenance (eg, program institutionalization). Both reviewers extracted data from all studies, and disagreements were resolved through discussion.

Two reviewers (A.P. and S.J.) assessed risk of bias using the Cochrane Collaboration’s tool for assessing risk of bias.^{36,37} Domains of bias included selection, performance, detection, attrition, reporting, and other. Reviewers independently evaluated risk of bias, assigning “low risk of bias,” “high risk of bias,” or “unclear risk of bias” to each domain. Discrepancies were resolved by consensus.

Results

Nine studies were included in the review.²¹⁻²⁹ The studies were randomized controlled trials,²¹⁻²⁸ with one cluster randomized controlled trial.²⁹ The majority of trials were conducted in a college or university setting ($n = 5$)^{21,24-27} or a clinical setting ($n = 3$),^{22,23,29} with one conducted in the community surrounding a university.²⁸ Seven studies were conducted in the United States,^{21,22,24-28} 1 study in the United Kingdom,²³ and 1 study in Finland.²⁹ In regard to intervention delivery, 3 studies included in-person interventions^{26,28,29} and 3 interventions solely utilized an online or other electronic platform.^{21,24,25} One study had the option of an in-person, online, or hybrid course.²⁷ Two studies utilized in-person intervention delivery with electronic communication follow-up.^{22,23} Additional details are presented in Table 1.

A summary of the RE-AIM results by each element is provided in Table 2; detailed results are available in Supplemental Digital Content Appendix Tables 2-4 (available at <http://links.lww.com/JPHMP/A650>). Harden et al³⁵ included a total of 60 RE-AIM recommended reporting criteria. Of these, 8 criteria were reported by all 9 studies, 26 criteria were reported by 4 or fewer studies, and 22 criteria were not reported by any of the studies. The 8 reported criteria were consistent with those required by current CONSORT guidelines (ie, inclusion/exclusion criteria, attrition, and number, timing, duration of contacts).

Reach was evaluated by 12 criteria including descriptions of who was intended to benefit (ie, the target population), who actually participated or was exposed to the intervention, how many persons participated out of those intended or targeted, and the characteristics of those who took part compared with those who did not.¹⁵ The target population was described by all studies but was most often limited in detail, including 2 characteristics (eg, female students, aged 18-30 years), making it difficult to determine a denominator for the percentage of the target population reached. Three studies only recruited women,^{26,28,29} while the other 6 studies included a majority of female participants (51%-78%).^{21-25,27} A majority of participants in all studies were white (62%-80%).²¹⁻²⁹ Three of the 9 studies referenced attempts to address diversity/representation in their participant pool.^{22,23,27} Participants in 4 studies had an average baseline BMI in the normal weight range,^{21,24,25,29} participants in 4 studies fell in the overweight range,^{22,26-28} and 1 study had participants in the obese range.²³ Three studies required a BMI above 18.5 kg/m² to avoid participants falling below normal

weight range.^{22,26,27} All studies reported recruitment strategies, inclusion/exclusion criteria for study participants, and sample size. Two studies described differences and similarities between the target and study populations.^{25,27} Only one study reported on participant eligibility and individual participation rate.²⁹ None of the studies reported on recommended reach criteria or statistical comparisons between the target and study populations, cost of recruitment, or use of qualitative methods to measure reach or participation rates.

Effectiveness (or efficacy) was evaluated by 9 criteria including the degree to which the intervention changes health outcomes and quality of life, taking into account unintended or negative results.¹⁵ Six studies found the intervention had no effect on BMI outcomes between the control and intervention groups at follow-up.^{23-25,27-29} Three studies found statistically significant differences in change in weight or BMI between the intervention and control groups.^{21,22,26} Kattelman et al²⁵ included gender as a fixed effect in their model to account for different retention rates between males and females. Three studies^{21,22,38} cited unintended consequences of the intervention including reduction below normal weight²⁴ and rapid weight change associated with self-weighing.²¹ All studies addressed attrition; none of the studies addressed cost-effectiveness or qualitative measures of effectiveness.

Adoption was assessed at the setting and individual provider levels (by 10 and 11 criteria, respectively) including the number and proportion of settings and staff members who agreed to participate in delivering the intervention and how representative they were of the intended audience in terms of the setting and staff.¹⁵ One study reported 8 of the 10 setting criteria, allowing for calculation of the setting-level participation rate and reporting of the average number of persons served per participating location.²⁷ Five studies reported on the number of participating sites,^{22,24,25,27,29} whereas 6 studies described intervention location.^{21,24-27,29} None of the studies reported on the criteria of comparisons between targeted and participating sites. With regard to adoption by providers, only 4 studies reported the level of expertise of the intervention providers^{22,26,28,29} or training^{28,29} and supervision of the intervention providers.^{26,28,29} The adoption or participation rate for individual providers was not calculable for any studies, since only one study reported the number of participating intervention staff or providers²⁸ and none reported the number of eligible individual providers or their characteristics. Similarly, none of the studies reported differences between targeted and participating providers

TABLE 1
Weight Gain Prevention Study Details

Author (Publication Year)	Baseline Characteristics, Age, Mean (SD), BMI, Mean (SD), Sample Size (n)	Setting, Intervention Delivery Method	Duration of Intervention, Follow-up	Weight Outcome, Change From Baseline to Follow up, kg ^a
Bertz et al (2015) ²¹	Female: 51% White: 64% Age: 19 (0.4) BMI (C/I): 23.0 (3.1) kg/m ² , 22.7 (2.9) kg/m ² n = 167, sites = 1	College campus Wi-Fi-enabled scales Web platform	I: 12 mo F: none	Mean (SD): C: 1.1 (4.4) I: -0.5 (3.7) P = .035
Biddle et al (2015) ²³	Female: 68.5% White: 80.2% Age: 32.8 (5.6) BMI: 34.6 (4.9) kg/m ² n = 187	UK primary care facilities In-person workshop Physical activity tracker Follow-up calls	I: 12 mo F: none	Mean (95% CI): C: -1.02 (-2.63, 0.58) I: -0.87 (-2.74, 1.00) P = .869
Greene et al (2012) ²⁴	Female: 63% White: 79% Age: 19.1 (1.1) BMI: 23.9 (4.1) kg/m ² n = 1689, sites = 8	College campuses Online platform, e-mail	I: 3 mo F: 15 mo	BMI, ^b mean (SE): C: 23.5 (0.19) - 23.9 (0.20) kg/m ² I: 23.3 (0.20) - 23.5 (0.21) kg/m ² P > .05
Kattelman et al (2014) ²⁵	Female: 67% White: 72.1% Age: 19.3 (1.1) BMI: 24.1 (4.4) kg/m ² n = 1639, sites = 1	College campus Online platform E-mail	I: 10 wk F: 15 mo	Mean (SD): C: 69.9 (16.2) - 70.6 (16.3) I: 68.6 (14) - 69.1 (13.8) P = .39
Katterman et al (2014) ²⁶	Female: 100% White: 62% Age, median (range): 22.3 (18-29) BMI: 26.63 kg/m ² n = 58, sites = 1	College campus In-person group meetings	I: 16 wk F: 12 mo	Estimated marginal means: C: +1.07 kg I: -2.24 kg P = .008
Lytle et al (2017) ²⁷	Female: 67.6% White: 72.6% Age: 22.7 (5.0) BMI: 25.4 (3.8) kg/m ² n = 441, sites = 3	Community college College course (in person, online, or hybrid)	I: 4 mo F: 24 mo	Mean (SD): C: 74.4 (0.863) I: 73.8 (0.857) P = .707
Metzgar and Nickols-Richardson (2016) ²⁸	Female: 100% White: 66% Age: 31.4 (8.1) BMI: 27.9 (6.8) kg/m ² n = 87	Community around college In-person group meetings	I: 12 mo F: none	Mean (SE): C: 77.9 (1.9) - 77.2 (2.2) RDs: 73.9 (1.6) - 75.2 (1.9) Counselor: 74.2 (1.1) - 75.1 (1.3) P > .05
Valve et al (2013) ²⁹	Female: 100% White: N/A Age, median (range): 19 (17-21) BMI: 22 (4.0) kg/m ² n = 1537, sites = 8	Vaccination centers (Finland) One-on-one counseling	I: 1.5-2.5 y F: none	BMI, ^b median (IQR): I: 0.55 (1.59) C: 0.51 (1.75) P = .996
Wing et al (2016) ²²	Female: 78% White: 73% Age: 28.2 (4.4) BMI: 25.4 (2.6) kg/m ² n = 599, sites = 2	Clinical sites In-person group meetings, online refresher course, email	I: 4 mo F: 3 y (average)	Mean (SE): C: 0.26 (0.22) kg SC: -0.56 (0.22) LC: -2.37 (0.22) P (C vs SC) = .018 P (C vs LC) < .001 P (S vs LC) < .001

Abbreviations: BMI, body mass index; C, control; I, intervention; IQR, interquartile range; F, follow-up; LC, large changes; N/A, not available; RD, registered dietitian; SC, small changes.

^aBoldface indicates statistical significance ($P < .05$).

^bBMI reported as primary outcome, see the Supplemental Digital Content Appendix (available at: <http://links.lww.com/JPHMP/A650>) for additional details.

TABLE 2
RE-AIM Criteria²⁸ Included in Each Study

RE-AIM Element	Criteria	Study										Total
		Bertz et al (2015) ²¹	Biddle et al (2015) ²³	Greene et al (2012) ²⁴	Kattelmann et al (2014) ²⁵	Katterman et al (2014) ²⁶	Lytle et al (2017) ²⁷	Metzgar and Nickols-Richardson (2016) ²⁸	Valve et al (2013) ²⁹	Wing et al (2016) ²²		
Reach	Described target population	X	X	X	X	X	X	X	X	X	X	9
	Demographic, behavioral information about target population		X		X							1
	Method to identify the target population	X	X	X	X	X	X	X	X	X	X	1
	Recruitment strategies	X	X	X	X	X	X	X	X	X	X	9
	Inclusion/exclusion criteria for individuals	X	X	X	X	X	X	X	X	X	X	9
	Eligible, invited (exposed to recruitment) potential participants	X	X	X	X	X	X	X	X	X	X	1
	Sample size	X	X	X	X	X	X	X	X	X	X	9
	Individual participation rate (sample size/eligible invited potential participants)											1
	Comparisons between the target population and the study sample				X		X					2
	Statistical comparisons between the target population and the study sample											0
Effectiveness	Cost of recruitment											0
	Qualitative methods to measure reach											0
	Report of mediators					X						1
	Report of moderators	X		X	X					X	X	4
	Intent-to-treat	X	X						X	X	X	5
	Imputation procedures		X				X			X	X	3
	Quality-of-life measures		X							X	X	2
	Unintended consequences measures/results	X	X	X	X					X	X	3
	Percent attrition (at program completion)	X	X	X	X	X	X	X	X	X	X	9
	Cost-effectiveness											0
Qualitative methods to measure efficacy/effectiveness											0	

(continues)

TABLE 2
RE-AIM Criteria²⁸ Included in Each Study (Continued)

RE-AIM Element	Criteria	Study										Total		
		Bertz et al (2015) ²¹	Biddle et al (2015) ²³	Greene et al (2012) ²⁴	Kattelmann et al (2014) ²⁵	Katterman et al (2014) ²⁶	Lytle et al (2017) ²⁷	Metzgar and Nickols-Richardson (2016) ²⁸	Valve et al (2013) ²⁹	Wing et al (2016) ²²				
Adoption, setting	Eligible, invited potential settings						X							1
	Number of participating settings			X	X		X					X		5
	Setting participation rate						X							1
	Description of the targeted location						X							1
	Inclusion/exclusion criteria of the setting						X							1
	Description of intervention location	X		X	X	X	X					X		6
	Method to identify the setting						X							1
	Comparisons between the targeted and participating settings													0
	Statistical comparisons between the targeted and participating settings													0
	Average number of persons served per setting										X			1
Adoption, provider	Eligible, invited potential providers (staff)													0
	Number of participating providers (staff)										X			1
	Provider (staff) participation rate													0
	Method to identify target providers													0
	Level of expertise of providers								X					4
	Inclusion/exclusion criteria for providers													0
	Comparisons between targeted and participating providers (staff)										X			0
	Statistical comparisons between targeted and participating providers (staff)													0
	Measures of cost adoption													0
	Dissemination beyond originally planned													0
Qualitative methods to measure adoption													0	

(continues)

or statistical comparisons between these groups, cost of adoption, dissemination of the intervention beyond where originally planned, or use of qualitative methods to measure individual provider adoption.

Implementation was assessed by the degree to which studies reported on 9 criteria including whether settings and staff members delivered the intervention as intended, the fidelity of intervention delivery, and costs.¹⁵ All studies reported individual participant engagement in terms of number, timing, duration of contact, and participant attendance. Only one study²² described whether the intervention protocol was delivered as intended, reporting that sessions sampled for measurement (20% of all sessions delivered) presented appropriate behavioral content. One study²⁸ reported qualitative data regarding quality of intervention content delivery between individual providers; another²⁷ included information about consistency of implementation across settings. Finally, one study²⁷ reported partial costs, that is, tuition costs of the for-credit course made available to intervention group participants free of charge.

Maintenance was evaluated as the sustained effectiveness at the participant level (per 3 criteria), and the sustained delivery of the intervention at the setting or staff level (per 6 criteria), including the alignment of the intervention with organizational mission, objectives, and goals and integration into job descriptions and performance evaluations.¹⁵ At participant level, 3 studies included follow-up outcome measures after intervention termination, of which all reported 60% to 70% retention at follow-up and some of which differed by race/ethnicity, age, gender, and baseline BMI of study participants.^{23,24,26} At the setting level, only one study reported attrition of intervention sites at follow-up, reporting no sites lost to follow-up.²⁷ None of the studies reported on any of the other criteria including qualitative methods to capture individual-level outcomes or maintenance of changes on whether the intervention was still in place after completion of the research study.

Risk of bias

Using the Cochrane risk of bias tool,³⁷ the 9 studies included in this review were overall rated low to unclear risk of bias (see Supplemental Digital Content Appendix Table 5, available at <http://links.lww.com/JPHMP/A650>). The greatest source of bias from the reviewed studies was performance bias, as blinding of participants and study personnel was often not conducive with the study design.^{22-27,29} One study²⁸ blinded study participants to group assignment and it was unclear whether a second study²¹ blinded

participants. Two studies had a high risk of attrition bias due to a large volume of missing follow-up data²³ and difference in BMI and desire to lose weight between completers and noncompleters.²⁴ We assessed low risk of bias due to cluster randomized study design for Valve et al²⁹; there was no recruitment bias, as clusters were randomized after recruitment and analysis was appropriate for a cluster design.³⁷

Discussion

The purpose of this review was to use RE-AIM to assess the extent to which weight gain prevention studies targeting young adults, and reporting results within the past decade, included elements of external validity. Our results suggest that there remains inadequate reporting on recent weight gain prevention studies with regard to components of external validity and generalizability.³⁹ Issues critical to translating research findings to public health impact often receive little attention when compared with intervention efficacy in narrow research settings.⁴⁰ This is a significant scientific constraint that limits the information required to disseminate and implement these interventions for population impact.^{39,41}

This review offers several insights into the comprehensiveness of reporting by studies on weight gain prevention. First, there is an overall general lack of reporting by studies on all RE-AIM criteria. Of the total of 60 RE-AIM criteria,³⁵ 37% (n = 22) were not reported by any studies and only 13% (n = 8) were reported by all studies. Despite the call for more comprehensive presentation of weight gain prevention study results, this dearth of reporting on elements of external validity shows that there has been minimal improvement in the past decade.^{9,10,42,43} The lack of information regarding external validity greatly limits interpretation and comparisons across studies that are required to fully understand impact and to inform future research efforts.^{15,44} Consistent reporting of external validity of weight gain prevention studies is needed to more effectively translate results into evidence-based policy and practice and to push the field to incorporate external validity into study planning and execution.

Second, the RE-AIM elements that were most often reported aligned with elements often required by journal or CONSORT publication guidelines.⁴⁵ For example, of the 15 criteria for reach, all studies reported on the 4 criteria required by CONSORT guidelines.⁴⁵ In contrast, 2 or fewer studies reported on the remaining 9 reach criteria such as enrollment, recruitment and participation rates, or costs of recruitment, which are rarely required for publication.^{38,46} Publication requirements appear to influence whether elements of

external validity are, or are not, addressed. Glasgow and colleagues¹⁶ have proposed an expanded CONSORT figure to increase the transparency in reporting external validity. Given the adherence to current guidelines, requiring this expanded guideline has the potential to enhance reporting.

It is also important to note that only 3 of 9 studies reported a significant effect on the primary outcome of weight but that there was minimal information provided on external factors needed to understand the full scope, or lack thereof, of intervention effectiveness. Without detailed reporting on setting, provider, and participant adoption and fidelity, it is not possible to determine whether interventions found not to impact weight were not successful due to the intervention itself or due to implementation failure or lack of engagement. Consistent and comprehensive reporting is needed to inform the science of what and how interventions work, and who they work best with, to improve the development of future interventions.

Finally, there remains a dearth of reporting on differences between settings and providers who accept or decline to adopt an intervention.¹⁶ This makes it difficult to determine what criteria might be needed for a site to successfully deliver the intervention, who in a real-world setting is best suited to deliver the intervention, or what settings might be appropriate for translation. There were also significant gaps in how implementation or maintenance was reported in these studies, including how consistently an intervention was delivered, whether adaptations to the original intervention were made, and elements of intervention continuation.⁴¹ This makes it difficult to determine whether a weight gain prevention intervention can be effectively delivered, in what setting, by whom it can be delivered, and whether it is sustainable.^{15,47} These reporting omissions prevent the timely dissemination of interventions and contribute to the decades-long gap between research and real-world practice.¹⁶

In summary, to enhance the impact of weight gain prevention intervention studies on a population level, the emphasis on designing and executing studies to produce generalizability findings and the reporting of external validity elements must improve. Over the past decade, support for transparency in research⁴⁸ has resulted in tools and checklists to aid in a balanced reporting process.^{13,14,45,49} Adoption of the expanded 2017 CONSORT criteria for nonpharmacologic clinical trials, which include both internal and external validity elements,^{14,45} also encourages consistent and balanced reporting.¹⁶ Standardizing requirements to include components of external validity, such as those proposed by Glasgow and colleagues,¹⁶ will improve

the quality, comprehensiveness, and consistency of study reporting, necessary for the better interpretation and understanding of findings of current studies. Utilization of tools such as RE-AIM⁴¹ and PRECIS-2¹⁷ to help design weight gain prevention studies with a greater focus on external validity in addition to consistent reporting of external validity components of studies is needed to more effectively translate results into evidence-based policy and practice.

Strengths and limitations

To our knowledge, there have been no other reviews of weight gain prevention interventions among young adults using the RE-AIM framework to address external validity. This review expands on recommendations from several prior reviews to address rigor and external validity of research related to annual weight gain as a critical obesity prevention target. In addition to careful abstraction of relevant studies by research staff, 2 expert reviewers further assessed studies using the RE-AIM tool. Limitations included the risk of bias due to study attrition and inability to compare outcomes across studies due to variation in reporting.

Conclusion

Prevention of weight gain in young adults is critical to reversing the obesity epidemic.⁴ Despite a heightened focus on balanced reporting of study validity, there remains inadequate reporting of prevention of weight gain studies with regard to elements of external validity and generalizability. The continued lack of prioritizing generalizability in study design and execution and reporting on dimensions of external validity is a significant scientific constraint that limits opportunities to disseminate and implement prevention

Implications for Policy & Practice

- **Practice:** Reporting on external validity is needed to determine whether a weight gain prevention intervention can be effectively delivered, in what setting, by whom it can be delivered, and whether it is sustainable in practice.
- **Policy:** Consistent reporting of external validity of weight gain prevention studies is needed to more effectively translate results into evidence-based policy and practice.
- **Research:** The lack of generalizable findings from studies designed to prioritize primarily internal validity and the lack of information regarding external validity greatly limit interpretation and comparisons across studies that are required to fully understand impact and to translate research to practice.

of weight gain interventions for population impact. Standardized reporting may be needed to ensure results that demonstrate not only internal validity¹⁶ but also external validity and generalizability are needed to promote public health impact.¹⁵

References

- Gordon-Larsen P, The NS, Adair LS. Longitudinal trends in obesity in the United States from adolescence to the third decade of life. *Obesity (Silver Spring)*. 2010;18(9):1801-1804.
- Lewis CE, Jacobs DR Jr, McCreath H, et al. Weight gain continues in the 1990s: 10-year trends in weight and overweight from the CARDIA study. Coronary Artery Risk Development in Young Adults. *Am J Epidemiol*. 2000;151(12):1172-1181.
- Woolf SH, Schoemaker H. Life expectancy and mortality rates in the United States, 1959-2017. *JAMA*. 2019;322(20):1996-2016.
- Dietz WH. Obesity and excessive weight gain in young adults: new targets for prevention. *JAMA*. 2017;318(3):241-242.
- Truesdale KP, Stevens J, Lewis CE, Schreiner PJ, Loria CM, Cai J. Changes in risk factors for cardiovascular disease by baseline weight status in young adults who maintain or gain weight over 15 years: the CARDIA study. *Int J Obes (Lond)*. 2006;30(9):1397-1407.
- Laska MN, Pelletier JE, Larson NI, Story M. Interventions for weight gain prevention during the transition to young adulthood: a review of the literature. *J Adolesc Health*. 2012;50(4):324-333.
- Hutfless S, Gudzone KA, Maruthur N, et al. Strategies to prevent weight gain in adults: a systematic review. *Am J Prev Med*. 2013;45(6):e41-e51.
- Shadish WR, Cook TD, Campbell DT. *Experimental and Quasi-experimental Designs for Generalized Causal Inference*. Boston, MA: Houghton Mifflin & Company; 2002.
- Willmott TJ, Pang B, Rundle-Thiele S, Badejo A. Weight management in young adults: systematic review of electronic health intervention components and outcomes. *J Med Internet Res*. 2019;21(2):e10265.
- Partridge SR, Juan SJ, McGeechan K, Bauman A, Allman-Farinelli M. Poor quality of external validity reporting limits generalizability of overweight and/or obesity lifestyle prevention interventions in young adults: a systematic review. *Obes Rev*. 2015;16(1):13-31.
- Glasgow RE, Emmons KM. How can we increase translation of research into practice? Types of evidence needed. *Annu Rev Public Health*. 2007;28:413-433.
- Huebschmann AG, Leavitt IM, Glasgow RE. Making health research matter: a call to increase attention to external validity. *Annu Rev Public Health*. 2019;40:45-63.
- Pinnock H, Barwick M, Carpenter CR, et al. Standards for reporting implementation studies (StaRI) statement. *BMJ*. 2017;356:i6795.
- Altman DG, Schulz KF, Moher D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med*. 2001;134(8):663-694.
- Glasgow RE, Estabrooks PE. Pragmatic applications of RE-AIM for health care initiatives in community and clinical settings. *Prev Chronic Dis*. 2018;15:E02.
- Glasgow RE, Huebschmann AG, Brownson RC. Expanding the CONSORT figure: increasing transparency in reporting on external validity. *Am J Prev Med*. 2018;55(3):422-430.
- Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. *BMJ*. 2015;350:h2147.
- Glasgow RE, Harden SM, Gaglio B, et al. RE-AIM planning and evaluation framework: adapting to new science and practice with a 20-year review. *Front Public Health*. 2019;7:64.
- RE-AIM. Welcome to RE-AIM.org. re-aim.org. Published 2019. Accessed February 2, 2019.
- Booth A, Clarke M, Dooley G, et al. The nuts and bolts of PROSPERO: an international prospective register of systematic reviews. *Syst Rev*. 2012;1(1):2.
- Bertz F, Pacanowski CR, Levitsky DA. Frequent self-weighing with electronic graphic feedback to prevent age-related weight gain in young adults. *Obesity (Silver Spring)*. 2015;23(10):2009-2014.
- Wing RR, Tate DF, Espeland MA, et al. Innovative self-regulation strategies to reduce weight gain in young adults: the Study of Novel Approaches to Weight Gain Prevention (SNAP) randomized clinical trial. *JAMA Intern Med*. 2016;176(6):755-762.
- Biddle SJ, Edwardson CL, Wilmot EG, et al. A randomised controlled trial to reduce sedentary time in young adults at risk of type 2 diabetes mellitus: Project STAND (Sedentary Time ANd Diabetes). *PLoS One*. 2015;10(12):e0143398.
- Greene GW, White AA, Hoerr SL, et al. Impact of an online healthful eating and physical activity program for college students. *Am J Health Promot*. 2012;27(2):e47-e58.
- Kattelman KK, Bredbenner CB, White AA, et al. The effects of Young Adults Eating and Active for Health (YEAH): a theory-based Web-delivered intervention. *J Nutr Educ Behav*. 2014;46(6):S27-S41.
- Katterman SN, Goldstein SP, Butryn ML, Forman EM, Lowe MR. Efficacy of an acceptance-based behavioral intervention for weight gain prevention in young adult women. *J Contextual Behav Sci*. 2014;3(1):45-50.
- Lytle LA, Laska MN, Linde JA, et al. Weight-gain reduction among 2-year college students: the CHOICES RCT. *Am J Prev Med*. 2017;52(2):183-191.
- Metzgar CJ, Nickols-Richardson SM. Effects of nutrition education on weight gain prevention: a randomized controlled trial. *Nutr J*. 2016;15:31.
- Valve P, Lehtinen-Jacks S, Eriksson T, et al. LINDA—a solution-focused low-intensity intervention aimed at improving health behaviors of young females: a cluster-randomized controlled trial. *BMC Public Health*. 2013;13:1044.
- Wilmot EG, Davies MJ, Edwardson CL, et al. Rationale and study design for a randomised controlled trial to reduce sedentary time in adults at risk of type 2 diabetes mellitus: project stand (Sedentary Time ANd diabetes). *BMC Public Health*. 2011;11:908.
- Moe SG, Lytle LA, Nanney MS, Linde JA, Laska MN. Recruiting and retaining young adults in a weight gain prevention trial: lessons learned from the CHOICES study. *Clin Trials*. 2016;13(2):205-213.
- Metzgar CJ, Nickols-Richardson SM. Determinants of weight gain prevention in young adult and midlife women: study design and protocol of a randomized controlled trial. *JMIR Res Protoc*. 2015;4(1):e36.
- Kattelman KK, White AA, Greene GW, et al. Development of Young Adults Eating and Active for Health (YEAH) Internet-based intervention via a community-based participatory research model. *J Nutr Educ Behav*. 2014;46(2):S10-S25.
- Covidence systematic review software, Veritas Health Innovation, Melbourne, Australia. www.covidence.org.
- Harden SM, Gaglio B, Shoup JA, et al. Fidelity to and comparative results across behavioral interventions evaluated through the RE-AIM framework: a systematic review. *Syst Rev*. 2015;4:155.
- Higgins JPT, Altman DG, Gotzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*. 2011;343:d5928.
- Higgins JPT, Green S, eds. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0*. London, England: The Cochrane Collaboration; 2011.
- Green LW, Glasgow RE. Evaluating the relevance, generalization, and applicability of research: issues in external validation and translation methodology. *Eval Health Prof*. 2006;29(1):126-153.
- Tomoaia-Cotisel A, Scammon DL, Waitzman NJ, et al. Context matters: the experience of 14 research teams in systematically reporting contextual factors important for practice change. *Ann Fam Med*. 2013;11(suppl 1):S115-S123.
- RE-AIM. Applying the RE-AIM framework. About RE-AIM. http://www.re-aim.org/about/applying-the-re-aim-framework. Accessed November 18, 2019.
- Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. *Am J Public Health*. 1999;89(9):1322-1327.
- Hebden L, Chey T, Allman-Farinelli M. Lifestyle intervention for preventing weight gain in young adults: a systematic review and meta-analysis of RCTs. *Obes Rev*. 2012;13(8):692-710.

43. Hayba N, Partridge S, Nour M, Grech A, Allman Farinelli M. Effectiveness of lifestyle interventions for preventing harmful weight gain among young adults from lower socioeconomic status and ethnically diverse backgrounds: a systematic review. *Obes Rev.* 2018;19(3):333-346.
44. Gaglio B, Shoup JA, Glasgow RE. The RE-AIM framework: a systematic review of use over time. *Am J Public Health.* 2013;103(6):e38-e46.
45. Boutron I, Altman DG, Moher D, Schulz KF, Ravaud P. CONSORT statement for randomized trials of nonpharmacologic treatments: a 2017 update and a CONSORT extension for nonpharmacologic trial abstracts. *Ann Intern Med.* 2017;167(1):40-47.
46. Glasgow RE, Askew S, Purcell P, et al. Use of RE-AIM to address health inequities: application in a low-income community health center-based weight loss and hypertension self-management program. *Transl Behav Med.* 2013;3(2):200-210.
47. McCrabb S, Lane C, Hall A, et al. Scaling-up evidence-based obesity interventions: a systematic review assessing intervention adaptations and effectiveness and quantifying the scale-up penalty. *Obes Rev.* 2019.
48. National Institutes of Health. *Reviewer Guidance on Rigor and Transparency: Research Project Grant and Mentored Career Development Applications.* Bethesda, MD: National Institutes of Health; 2019.
49. Hoffmann TC, Glasziou PP, Boutron I, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ.* 2014;348:g1687.