Outcomes addressed in randomized controlled lifestyle intervention trials in community-dwelling older people with (sarcopenic) obesity—An evidence map

Isabel Galicia Ernst¹ | Gabriel Torbahn¹,² | Lukas Schwingshackl³ | Helge Knüttel⁴ | Robert Kob¹ | Wolfgang Kemmler⁵,⁶ | Cornel C. Sieber¹,⁷ | John A. Batsis⁸ | Dennis T. Villareal⁹ | Nanette Stroebele-Benschop¹⁰ | Marjolein Visser¹¹ | Dorothee Volkert¹ | Eva Kiesswetter¹,³ | Daniel Schoene⁵,¹²

¹Institute for Biomedicine of Aging, Friedrich-Alexander-Universität Erlangen-Nürnberg, Nuremberg, Germany
²Department of Pediatrics, Paracelsus Medical University, Klinikum Nürnberg, Universitätsklinik der Paracelsus Medizinischen Privatuniversität Nürnberg, Nuremberg, Germany
³Institute for Evidence in Medicine, Faculty of Medicine and Medical Center, University of Freiburg, Freiburg, Germany
⁴University Library, University of Regensburg, Regensburg, Germany
⁵Institute of Medical Physics, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany
⁶Institute of Radiology, University Hospital Erlangen, Erlangen, Germany
⁷Department of Medicine, Kantonsspital Winterthur, Winterthur, Switzerland
⁸Division of Geriatric Medicine, School of Medicine and Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA
⁹Division of Endocrinology, Diabetes and Metabolism, Baylor College of Medicine, Houston, Texas, USA
¹⁰Department of Nutritional Psychology, Institute of Nutritional Medicine, University of Hohenheim, Stuttgart, Germany
¹¹Department of Health Sciences, Faculty of Science, Vrije Universiteit Amsterdam, Amsterdam Public Health Research Institute, Amsterdam, The Netherlands
¹²Institute for Exercise and Public Health, University of Leipzig, Leipzig, Germany

Correspondence
Isabel Galicia Ernst, MSc, Institute for Biomedicine of Aging, Friedrich-Alexander-Universität Erlangen-Nürnberg, Nuremberg, Bavaria, Germany, Kobergerstrasse 60, 90408, Nuremberg, Bavaria, Germany.
Email: isabel.galicia.ernst@fau.de

Summary
Obesity and sarcopenic obesity (SO) are characterized by excess body fat with or without low muscle mass affecting bio-psycho-social health, functioning, and subsequently quality of life in older adults. We mapped outcomes addressed in randomized controlled trials (RCTs) on lifestyle interventions in community-dwelling older people with (sarcopenic) obesity. Systematic searches in Medline, Embase, Cochrane Central, CINAHL, PsycInfo, Web of Science were conducted. Two reviewers independently performed screening and extracted data on outcomes, outcome domains, assessment

Abbreviations: RCTs, randomized controlled trials; SO, sarcopenic obesity; COS, core outcome set; BMI, body mass index; COA, clinical outcome assessment; FDA, food and drug administration; PRO, patient-reported outcome; ORO, observer-reported outcome; CRO, clinician-reported outcome; PerFO, performance outcome; BM, biomarkers; VO₂max/peak, peak oxygen uptake; (I)ADLs, instrumental activities of daily living; SPPB, short physical performance battery; DXA, dual energy X-ray absorptiometry; BIA, bioelectrical impedance analysis; CT, computed tomography scan; MRI, magnetic resonance imaging; RNA, ribonucleic acid; HDL, high density lipoprotein; LDL, low density lipoprotein.

Eva Kiesswetter and Daniel Schoene equally shared last authorship.

Received: 14 March 2022 Revised: 27 June 2022 Accepted: 28 June 2022
DOI: 10.1111/obr.13497

© 2022 The Authors. Obesity Reviews published by John Wiley & Sons Ltd on behalf of World Obesity Federation.
1 | INTRODUCTION

Obesity and sarcopenic obesity (SO) in older adults are major public health issues, due to their increasing prevalence. Obesity prevalence among older adults (60 years and older) ranges from 20.9% in Europe to 43.3% in the United States. The prevalence of SO is difficult to establish due to the different definitions and cutoffs applied. Nevertheless, a recent meta-analysis from Gao et al. (2021) estimates that the global prevalence of SO for older adults (60 years and older) is estimated to be 11%. Both disorders are associated with negative health consequences, such as premature mortality, increased risk of falls, poor physical functioning, comorbidity burden, and reduced quality of life, increasing the risk of loss of independence and institutionalization.

The recommended first-line therapies for (sarcopenic) obesity in older adults are lifestyle interventions aiming at the loss of body weight and fat mass and consist of diet modifications, increasing physical activity or specific exercise training, and behavioural therapy. Reviews and guidelines on obesity treatment in older adults have been published suggesting that lifestyle interventions in older adults are effective in reducing body weight and favoring combined interventions including dietary and exercise components. The optimal content and dose of interventions, are not well established, one reason being the lack of systematic reviews with meta-analyses enabling to identify optimal treatment strategies. Clinical decisions about treatment should be based on outcomes of high-quality randomized controlled trials. Therefore, the selection of outcomes is important relative to adequate power, the fit with the target population, and the avoidance of null findings. The lack of pooled analyses might be due to the heterogeneity in assessed outcome domains, outcomes, units of measurement, and the time of outcome measurement. For better comparability of studies and purposes of data pooling, a harmonization of at least a core outcome set (COS) of important measures should exist. A COS helps avoid ineffective interventions and outcome-reporting bias by providing a list of the minimum outcomes to be measured in RCTs.

The development of a COS requires as a first step a comprehensive review of the existing literature and the extraction of the outcomes assessed as well as outcome-related methodology used in available RCTs. For this purpose, an evidence map is considered appropriate, listing the evidence, identifying gaps, and providing results in a user-friendly format. We created such an evidence map to provide an overview of outcomes reported in RCTs on lifestyle interventions in community-dwelling older people with (sarcopenic) obesity. Specifically, we addressed the following questions: Which outcomes from which domains have been measured in lifestyle intervention RCTs in community-dwelling older adults with (sarcopenic) obesity? Which methods were used and at which time points were the outcomes assessed? Do the identified outcomes and methods to assess these outcomes depend on whether obesity is treated or SO is treated?

2 | METHODS

This evidence map has been developed in the frame of the Effective SLOPE project (EffectS of Lifestyle interventions in Older PeoPle with obesity: a systematic review and network meta-analysis; PROSPERO: CRD42019147286). The reporting of this study complies with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) for Scoping Reviews (PRISMA-ScR) and the extension to the PRISMA Statement for Reporting Literature Searches in Systematic Reviews (PRISMA-S) (Tables S1 and S2).

2.1 | Information sources and search strategy

Briefly, six electronic databases (Medline [via Ovid], Embase [via Ovid], Cochrane Central Register of Controlled Trials [CENTRAL, via Cochrane Library], Cumulated Index to Nursing and Allied Health Literature [CINAHL, via EBSCOhost], PsycInfo [via EBSCOhost], Science Citation Index Expanded [SCI-EXPANDED, via Web of Science Core Collection/
Clarivate) and one trial registry (ClinicalTrials.gov) were searched for published, unpublished, or ongoing trials from inception or availability to the present. For the development of the search strategy, we used the search strategy from a Cochrane review on the effects of lifestyle interventions in children with obesity as a starting point. In addition, we used the Cochrane sensitivity-precision maximizing search filter for RCTs. For the search in CENTRAL, we did not consider entries from trial registries due to resource limitations. Search strategies were peer reviewed by an external information specialist (EM, University of Freiburg, Germany). Searches were fully re-run with the last search date in May 2022. When re-running the searches, records known from earlier searches (i.e., duplicates within a database) were removed based on their database accession numbers. We did not set any restrictions regarding language or publication time. The detailed search strategies are shown in supporting information Table S3. We further screened references lists of included reports for potentially relevant studies. Duplicates between databases were identified according to the method of Bramer et al. followed by a duplicate check in Covidence. Results from ClinicalTrials.gov were exported as CSV files and were screened according to eligibility criteria.

2.2 | Selection process

At least two reviewers (IGE, EK, DS, and GT) independently screened titles/abstracts and full texts according to prespecified eligibility criteria using Covidence systematic review software (Veritas Health Innovation, Melbourne, Australia; available at https://www.covidence.org). If information was lacking, we contacted the corresponding author/s two times over a span of 2 weeks. For translation of articles published in English, Spanish, or German, we involved colleagues who were fluent/native in these respective languages; for translation of articles published in other languages, we used online translators (e.g., https://www.deepl.com/home). Conference abstracts and theses not additionally published in a peer-reviewed journal were excluded.

The PRISMA 2020 compliant flow chart was created with the PRISMA template available on https://prisma-statement.org/prismastatement/flowdiagram.aspx.

2.3 | Eligibility criteria for RCTs

2.3.1 | Design of primary studies

We included RCTs (parallel and cluster).

2.3.2 | Population

Studies in community-dwelling adults with a minimum age of 60 years and a mean age of ≥65 years were considered. At least one obesity criterion for participants to be included in the RCTs had to be applied: total body fat mass ≥35% (women) and ≥25% (men), waist circumference of ≥88 cm (women) and ≥102 cm (men), and BMI using the standard adult cutoff of ≥30 kg/m². Cutoffs validated in specific populations (e.g., Asian [≥25 kg/m²]) were also accepted. Studies stating “obese” without providing a clear definition or criteria with references were excluded.

For SO, we used the definition provided in the primary RCTs, as long as one of the above-mentioned obesity criteria was met. RCTs in mixed samples of people with overweight and obesity were excluded. However, authors were contacted and asked for data on the obesity subgroup. If additional data were provided, the respective RCT was considered for the current analysis.

2.3.3 | Interventions

Lifestyle interventions were considered if the intervention consisted of diet modifications (e.g., calorie restriction), exercise (aerobic exercise, resistance exercise, or both), or behavioural therapy, as well as their combinations with all types of delivery and doses. Due to the time needed to respond to treatment, the minimum intervention duration was set to 12 weeks.

RCTs focusing only on very low energy diets (<800 kcal/day), total diet replacement, micronutrient supplements (e.g., vitamin D), secondary plant products (e.g., polyphenols), components of macronutrients (e.g., fatty acids [docosahexaenoic acid]), or amino acids (e.g., leucine), and dietary fibers were excluded.

2.3.4 | Comparators

As comparators, any lifestyle intervention and control groups (e.g., usual care and health counseling/education) were considered as a relevant comparator group.

2.3.5 | Outcomes

All reported health outcomes were deemed relevant. Related outcomes, such as environmental factors (e.g., walkability) and behaviour changes (e.g., level of physical activity and dietary intake) were not considered. Articles presenting the data on genetic outcomes only were excluded.

2.4 | Data extraction

Two reviewers (DS and IGE) extracted the data of included references independently (using a pre-piloted data extraction table in Microsoft Excel 2016). Disagreements were solved by discussion or with the help of a third reviewer (EK). For each RCT, the study characteristics (first author, year of publication, country, obesity phenotype, obesity criterion for inclusion, sample size, study duration, and mean age) were extracted. Further, to map relevant information on the outcomes assessed in the RCTs, a classification scheme (Table 1) was created and
used to extract (in addition to the outcome itself) data on outcome domain, type, (sub) category, method, units, and time of measurement.

2.5 | Data synthesis

Results on outcomes are presented for all included RCTs and separately for obesity and SO trials. To obtain an overview of the frequency of assessed outcomes for all included outcome domains, a bubble chart with four dimensions for each outcome domain was created. The x axis presents the four time-based categories of measurement (12 to 19 weeks, 20 to 26 weeks, 27 to 52 weeks, and more than 52 weeks). The y axis represents the number of studies reporting at least one outcome in each outcome domain. The size of the bubbles represents the number of outcomes for each domain, and the color of the bubbles represents the obesity phenotype addressed in the RCTs (obesity or SO).

Outcomes were counted and heatmaps were created for each domain based on the classification scheme in Table 1. A heatmap visualizes data in a compact form by representing numbers with corresponding colors. All outcomes assessed in at least two RCTs were counted and presented.

### TABLE 1 | Outcome classification scheme

<table>
<thead>
<tr>
<th>Classification criteria</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Outcome domains</strong></td>
<td>Grouping of outcomes into 10 generic domains according to the aim of measurement:</td>
</tr>
<tr>
<td></td>
<td>• Physical function</td>
</tr>
<tr>
<td></td>
<td>• Body composition and anthropometry</td>
</tr>
<tr>
<td></td>
<td>• Biomarkers</td>
</tr>
<tr>
<td></td>
<td>• Physiological</td>
</tr>
<tr>
<td></td>
<td>• Psychological</td>
</tr>
<tr>
<td></td>
<td>• Quality of life</td>
</tr>
<tr>
<td></td>
<td>• Sleep</td>
</tr>
<tr>
<td></td>
<td>• Pain</td>
</tr>
<tr>
<td></td>
<td>• Medications</td>
</tr>
<tr>
<td></td>
<td>• Risk for adverse health event and medical conditions</td>
</tr>
<tr>
<td><strong>Outcome type</strong></td>
<td>Defined according to the Clinical Outcome Assessment (COA) of the Food and Drug Administration (FDA)</td>
</tr>
<tr>
<td></td>
<td>• Patient-reported outcome (PRO): measurement based on a report that comes directly from the patients about their health status/condition.</td>
</tr>
<tr>
<td></td>
<td>• Observer-reported outcome (ORO): measurement based on a report of observable signs, events, or behaviors to a patient’s health condition by someone other than the patient or a health professional.</td>
</tr>
<tr>
<td></td>
<td>• Clinician-reported outcome (CRO): measurement based on a report that comes from a trained healthcare professional after observation of a patient’s health condition.</td>
</tr>
<tr>
<td></td>
<td>• Performance outcome (PerfO): measurement based on standardized tasks actively undertaken by a patient according to a set of instructions.</td>
</tr>
<tr>
<td></td>
<td>• Biomarkers (BM): a measurement that is considered as an indicator of normal biological processes, pathogenic processes, or biological responses to an exposure or intervention. May include molecular, histologic, radiographic, or physiologic characteristics.</td>
</tr>
<tr>
<td><strong>Outcome category and subcategory</strong></td>
<td>Specification of outcome domains in categories and subcategories (for psychological outcomes and QoL) based on the measurement aim, e.g., body composition was categorized in fat, muscle, and bone; biomarkers were categorized in glucose metabolism, lipids, hormones, and so on.</td>
</tr>
<tr>
<td><strong>Method</strong></td>
<td>Subsuming of methodological approaches (e.g., self-reported questionnaire), the used device (e.g., DXA), or sampling (e.g., blood) to assess the outcomes.</td>
</tr>
<tr>
<td><strong>Outcome</strong></td>
<td>• Health-related patient assessment used as an endpoint and providing a rating score (categorical or continuous).</td>
</tr>
<tr>
<td></td>
<td>• Outcomes with different names or slight differences in administration but addressing the same concept and measurement aim were categorized as one outcome (e.g., gait speed measured in 400 m or 10 m).</td>
</tr>
<tr>
<td></td>
<td>• Unique outcomes (outcomes in one RCT only) were reported separately. (supporting information Tables S5–S7).</td>
</tr>
<tr>
<td><strong>Unit</strong></td>
<td>Scoring and reported units of each outcome were listed.</td>
</tr>
<tr>
<td><strong>Time of measurement</strong></td>
<td>Based on information on baseline, intermediate/interim, post-intervention, and follow-up outcome assessment assigning of time measurement to one of the four categories:</td>
</tr>
<tr>
<td></td>
<td>1. 12–19 weeks</td>
</tr>
<tr>
<td></td>
<td>2. 20–26 weeks</td>
</tr>
<tr>
<td></td>
<td>3. 27–52 weeks</td>
</tr>
<tr>
<td></td>
<td>4. More than 52 weeks</td>
</tr>
<tr>
<td></td>
<td>When outcomes were measured more than once, all times were extracted.</td>
</tr>
</tbody>
</table>

Note: This table was the basis for creating the heatmaps. This information was extracted from each included randomized controlled trial (RCT).
considered for the heatmaps. The outcomes in the heatmaps were sorted based on their category. The heatmaps’ shade colors represent the frequency of outcome measurement. The shade colors were chosen according to traffic light colors, where green represents the outcomes most frequently reported, yellow represents a midpoint, and red the least reported outcomes.

All figures were created with the statistical software R Version 4.1.0. The bubble chart was generated using the ggplot2 package (v3.3.3 Wickham, New-York, USA, 2016). The heatmaps were created with the gt package (v0.3.0 Iannone, Boston, USA, 2021).

Separate heatmaps for each obesity phenotype were created for the domains physical function, body composition and anthropometry, and biomarkers. Due to a limited amount of data, only a combined heatmap, including both obesity and SO, was created for the domains quality of life, psychological, physiological, pain, sleep, medications, and risk for adverse health event and medical conditions.

### 3 | RESULTS

Of 57,721 unique records, 109 articles of 54 studies were included in the evidence map as they provided information on outcomes assessed (Figure 1 and Table S4). The unique accession numbers from all database searches are available upon request.

#### 3.1 | Study and participants’ characteristics

Table 2 presents the study and participants’ characteristics of included trials. Forty-seven (87.0%) RCTs included participants with obesity and seven (13.0%) participants with SO. Two RCTs defined SO using both muscle mass and muscle function, while five considered only muscle mass as criterion. The sample size ranged from 16 to 742 participants. Regarding study duration, 29 (53.7%) RCTs lasted between 12 and 19 weeks, 19 (35.2%) between 20 and 26 weeks, 2 (3.7%) 52 weeks, and 4 (7.4%) more than 52 weeks. No study in individuals with SO lasted longer than 26 weeks. Seven studies, all in individuals of the obesity phenotype, included a post-intervention follow-up period ranging from 5 to 19 months for at least some study outcomes. The study’s participants’ age ranged from 65.3 to 77.4 years across studies, with only four RCTs reporting a mean age of 75 years and older.

#### 3.2 | Outcome domains and outcomes

Nearly all studies \((n = 52, 96.3\%)\) reported body composition and anthropometry outcomes. Physical function \((n = 42, 77.8\%)\) and biomarkers \((n = 40, 74.1\%)\) outcomes were assessed in about three quarters of the trials. Twenty-three trials (42.6%) reported physiological outcomes, while outcomes from all other domains were less frequently reported: quality
<table>
<thead>
<tr>
<th>First author, year, country</th>
<th>Phenotype (obesity or sarcopenic obesity)</th>
<th>Obesity criterion for inclusion</th>
<th>Sample size</th>
<th>Mean age (years)</th>
<th>Study duration (weeks)</th>
<th>Outcome domains reported per RCT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abdelbasset 2020, Saudi Arabia</td>
<td>Obesity</td>
<td>BMI ≥ 30 kg/m²</td>
<td>40</td>
<td>71.3</td>
<td>12</td>
<td>PF, BC, PH</td>
</tr>
<tr>
<td>Amamou 2017, Canada</td>
<td>Obesity</td>
<td>WC &gt; 102 cm and &gt;88 cm</td>
<td>31</td>
<td>65.8</td>
<td>16</td>
<td>BC, BM, PH</td>
</tr>
<tr>
<td>Ard 2017, USA</td>
<td>Obesity</td>
<td>BMI 30–40 kg/m²</td>
<td>164</td>
<td>70.3</td>
<td>52</td>
<td>BC, BM, QoL, PF, PH</td>
</tr>
<tr>
<td>Balachandran 2014, USA</td>
<td>Sarcopenic obesity</td>
<td>BMI ≥ 30 kg/m²</td>
<td>21</td>
<td>71.3</td>
<td>15</td>
<td>PF, BC</td>
</tr>
<tr>
<td>Beavers 2015, USA</td>
<td>Obesity</td>
<td>BMI ≥ 27, WC ≥ 102 (men) and ≥88 (women)</td>
<td>25</td>
<td>68.4</td>
<td>12</td>
<td>PF, BC, BM, PH</td>
</tr>
<tr>
<td>Beavers 2019, USA</td>
<td>Obesity</td>
<td>BMI 30–40 kg/m²</td>
<td>96</td>
<td>70.3</td>
<td>26</td>
<td>PF, BC, BM, PS</td>
</tr>
<tr>
<td>Brennan 2020, USA</td>
<td>Obesity</td>
<td>BMI ≥ 30 kg/m²</td>
<td>86</td>
<td>68.6</td>
<td>26</td>
<td>PF, BC, BM, PH</td>
</tr>
<tr>
<td>Cai 2019, China</td>
<td>Obesity</td>
<td>BMI ≥ 28 kg/m²</td>
<td>480</td>
<td>66.8</td>
<td>104</td>
<td>BC, BM, PH</td>
</tr>
<tr>
<td>Davidson 2009, Canada</td>
<td>Obesity</td>
<td>WC &gt; 102 cm (men) and &gt;88 cm (women)</td>
<td>136</td>
<td>67.6</td>
<td>26</td>
<td>PF, BC, BM</td>
</tr>
<tr>
<td>Elsayed 2022, Egypt</td>
<td>Obesity</td>
<td>BMI 30–39.9 kg/m²</td>
<td>68</td>
<td>65.3</td>
<td>12</td>
<td>PF, BC, BM, PS</td>
</tr>
<tr>
<td>Fanning 2020, USA</td>
<td>Obesity</td>
<td>BMI 30–45 kg/m²</td>
<td>28</td>
<td>71.8</td>
<td>12</td>
<td>PF, BC, BM, PS, Pain</td>
</tr>
<tr>
<td>Fanning 2021, USA</td>
<td>Obesity</td>
<td>BMI 30–45 kg/m²</td>
<td>183</td>
<td>70.0</td>
<td>72</td>
<td>BC</td>
</tr>
<tr>
<td>Haywood 2017, Australia</td>
<td>Obesity</td>
<td>BMI ≥ 32 kg/m²</td>
<td>117</td>
<td>70.0</td>
<td>12</td>
<td>PF, BC, BM, PH, Medications</td>
</tr>
<tr>
<td>Kallings 2009, Sweden</td>
<td>Obesity</td>
<td>BMI ≥ 30 kg/m²</td>
<td>101</td>
<td>68.0</td>
<td>26</td>
<td>BC, BM, PH, QoL</td>
</tr>
<tr>
<td>Kemmler 2016, Germany</td>
<td>Sarcopenic obesity</td>
<td>&gt;35% body fat</td>
<td>75</td>
<td>77.0</td>
<td>26</td>
<td>PF, BC, BM, PH, R</td>
</tr>
<tr>
<td>Kemmler 2017, Germany</td>
<td>Sarcopenic obesity</td>
<td>&gt;27% body fat</td>
<td>100</td>
<td>77.4</td>
<td>16</td>
<td>PF, BC, BM, R</td>
</tr>
<tr>
<td>Kim 2018, South Korea</td>
<td>Obesity</td>
<td>BMI ≥ 30 kg/m²</td>
<td>20</td>
<td>66.4</td>
<td>12</td>
<td>BC, BM, PH</td>
</tr>
<tr>
<td>Kim 2019, South Korea</td>
<td>Obesity</td>
<td>≥25 kg/m² BMI, &gt;30% body fat</td>
<td>24</td>
<td>68.8</td>
<td>12</td>
<td>PF, BC, BM</td>
</tr>
<tr>
<td>Kim 2020, South Korea</td>
<td>Obesity</td>
<td>WC &gt; 90 cm (men) and &gt;85 cm (women)</td>
<td>75</td>
<td>74.9</td>
<td>12</td>
<td>PF, BC, QoL, BM</td>
</tr>
<tr>
<td>Kitzman 2016, USA</td>
<td>Obesity</td>
<td>BMI ≥ 30 kg/m²</td>
<td>100</td>
<td>67.0</td>
<td>20</td>
<td>PF, BC, QoL, BM, PH</td>
</tr>
<tr>
<td>Kritchevsky 2017, USA</td>
<td>Obesity</td>
<td>BMI ≥ 30 kg/m²</td>
<td>1176</td>
<td>77.1</td>
<td>140</td>
<td>PF, R</td>
</tr>
<tr>
<td>Lambert 2008, USA</td>
<td>Obesity</td>
<td>BMI ≥ 30 kg/m²</td>
<td>16</td>
<td>69.0</td>
<td>12</td>
<td>PF, BC, BM</td>
</tr>
<tr>
<td>Lee 2021, Taiwan</td>
<td>Sarcopenic obesity</td>
<td>Body fat &gt;35%</td>
<td>27</td>
<td>70.9</td>
<td>12</td>
<td>PF, BC</td>
</tr>
<tr>
<td>Maillard 2016, France</td>
<td>Sarcopenic obesity</td>
<td>BMI ≥ 25 kg/m² and ≤ 40 kg/m²</td>
<td>17</td>
<td>69.0</td>
<td>16</td>
<td>BC, BM</td>
</tr>
<tr>
<td>Manini 2010, USA</td>
<td>Obesity</td>
<td>BMI ≥ 30 kg/m²</td>
<td>179</td>
<td>75.6</td>
<td>62</td>
<td>PF, BC</td>
</tr>
<tr>
<td>Miller 2006, USA</td>
<td>Obesity</td>
<td>BMI ≥ 30.0 kg/m²</td>
<td>87</td>
<td>69.5</td>
<td>26</td>
<td>PF, BC, BM, Pain</td>
</tr>
<tr>
<td>Muscariello 2016, Italy</td>
<td>Sarcopenic obesity</td>
<td>BMI ≥ 30.0 kg/m²</td>
<td>104</td>
<td>66.7</td>
<td>12</td>
<td>PF, BC</td>
</tr>
<tr>
<td>Nabuco 2019, Brazil</td>
<td>Sarcopenic obesity</td>
<td>Body fat ≥ 35%</td>
<td>26</td>
<td>69.1</td>
<td>16</td>
<td>PF, BC, BM, PH</td>
</tr>
<tr>
<td>Nicklas 2014, USA</td>
<td>Obesity</td>
<td>BMI 30–40 kg/m²</td>
<td>48</td>
<td>70.1</td>
<td>20</td>
<td>BC</td>
</tr>
<tr>
<td>Nicklas 2019, USA</td>
<td>Obesity</td>
<td>BMI 30–45 kg/m²</td>
<td>180</td>
<td>69.2</td>
<td>20</td>
<td>PF, BC, QoL, BM, PS, PH</td>
</tr>
<tr>
<td>Normandin 2018, USA</td>
<td>Obesity</td>
<td>BMI 30–40 kg/m²</td>
<td>37</td>
<td>70.1</td>
<td>22</td>
<td>PF, BC, BM</td>
</tr>
<tr>
<td>O'Leary 2007, USA</td>
<td>Obesity</td>
<td>BMI 30–40 kg/m²</td>
<td>21</td>
<td>66.3</td>
<td>12</td>
<td>BC, BM</td>
</tr>
<tr>
<td>Park 2017, South Korea</td>
<td>Sarcopenic obesity</td>
<td>BMI ≥ 25 kg/m²</td>
<td>50</td>
<td>74.1</td>
<td>24</td>
<td>PF, BC, BM, PH</td>
</tr>
<tr>
<td>Park 2019, South Korea</td>
<td>Obesity</td>
<td>BMI ≥ 25 kg/m2</td>
<td>24</td>
<td>66.5</td>
<td>12</td>
<td>PF, BC, BM, PH</td>
</tr>
<tr>
<td>Park 2020, South Korea</td>
<td>Obesity</td>
<td>BMI ≥ 25 kg/m²</td>
<td>24</td>
<td>68.8</td>
<td>12</td>
<td>PF, BC, BM, PH</td>
</tr>
<tr>
<td>Porter Starr 2016, USA</td>
<td>Obesity</td>
<td>BMI ≥ 30 kg/m²</td>
<td>67</td>
<td>68.2</td>
<td>26</td>
<td>PF, BC, BM</td>
</tr>
</tbody>
</table>
TABLE 2 (Continued)

<table>
<thead>
<tr>
<th>First author, year, country</th>
<th>Phenotype (obesity or sarcopenic obesity)</th>
<th>Obesity criterion for inclusion</th>
<th>Sample size</th>
<th>Mean age (years)</th>
<th>Study duration (weeks)</th>
<th>Outcome domains reported per RCT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prieto 2014, Spain71</td>
<td>Obesity</td>
<td>BMI &gt; 30 kg/m²</td>
<td>76</td>
<td>67.1</td>
<td>24</td>
<td>PF, BC</td>
</tr>
<tr>
<td>Prieto 2015, Spain72</td>
<td>Obesity</td>
<td>BMI &gt; 30 kg/m²</td>
<td>56</td>
<td>67.2</td>
<td>24</td>
<td>PF, QoL</td>
</tr>
<tr>
<td>Rezaei-pour 2021, Iran73</td>
<td>Obesity</td>
<td>BMI &gt; 30 kg/m²</td>
<td>55</td>
<td>68.7</td>
<td>12</td>
<td>BC, BM</td>
</tr>
<tr>
<td>Rosenberg 2020, USA74</td>
<td>Obesity</td>
<td>BMI &gt; 30 kg/m²</td>
<td>60</td>
<td>68.0</td>
<td>12</td>
<td>PF, BC, BM, PH, QoL, PS, Sleep, Pain</td>
</tr>
<tr>
<td>Rosety 2015, Spain75</td>
<td>Obesity</td>
<td>BMI &gt; 30 kg/m²</td>
<td>48</td>
<td>67.7</td>
<td>12</td>
<td>PF, BC</td>
</tr>
<tr>
<td>Serra-Prat 2021, Spain76</td>
<td>Obesity</td>
<td>BMI &gt; 30 kg/m²</td>
<td>305</td>
<td>69.7</td>
<td>26</td>
<td>PF, BC, BM, QoL</td>
</tr>
<tr>
<td>Shah 2009, USA77</td>
<td>Obesity</td>
<td>BMI &gt; 30 kg/m²</td>
<td>18</td>
<td>68.6</td>
<td>24</td>
<td>PF, BC, BM, PH</td>
</tr>
<tr>
<td>Solomon 2009, USA78</td>
<td>Obesity</td>
<td>BMI &gt; 30 kg/m²</td>
<td>16</td>
<td>66.0</td>
<td>12</td>
<td>PF, BC</td>
</tr>
<tr>
<td>Solomon 2010, USA79</td>
<td>Obesity</td>
<td>BMI &gt; 30 kg/m²</td>
<td>22</td>
<td>66.0</td>
<td>12</td>
<td>PF, BC, BM, PH, R</td>
</tr>
<tr>
<td>Stillman 2018, USA80</td>
<td>Obesity</td>
<td>BMI &gt; 30 kg/m²</td>
<td>28</td>
<td>69.4</td>
<td>24</td>
<td>PF, BC, PS, QoL</td>
</tr>
<tr>
<td>Villareal 2006, USA81</td>
<td>Obesity</td>
<td>BMI &gt; 30 kg/m²</td>
<td>27</td>
<td>70.0</td>
<td>26</td>
<td>PF, BC, QoL, BM, PH</td>
</tr>
<tr>
<td>Villareal 2011, USA82</td>
<td>Obesity</td>
<td>BMI &gt; 30 kg/m²</td>
<td>107</td>
<td>69.7</td>
<td>52</td>
<td>PF, BC, QoL, BM, PH</td>
</tr>
<tr>
<td>Villareal 2017, USA83</td>
<td>Obesity</td>
<td>BMI &gt; 30 kg/m²</td>
<td>160</td>
<td>70.0</td>
<td>26</td>
<td>PF, BC, QoL, BM</td>
</tr>
<tr>
<td>Vincent 2014, USA84</td>
<td>Obesity</td>
<td>WC ≥ 102 cm for men, ≥ 88 cm for women, BMI &gt; 30 kg/m²</td>
<td>60</td>
<td>68.3</td>
<td>16</td>
<td>PF, BC, PS, PH, Medications, Pain</td>
</tr>
<tr>
<td>West 2011, USA85</td>
<td>Obesity</td>
<td>BMI &gt; 30</td>
<td>228</td>
<td>71.2</td>
<td>16</td>
<td>BC, PS</td>
</tr>
<tr>
<td>Yassine 2009, USA86</td>
<td>Obesity</td>
<td>BMI 30–40 kg/m²</td>
<td>24</td>
<td>65.5</td>
<td>12</td>
<td>PF, BC, BM, PH</td>
</tr>
<tr>
<td>Zhou 2021, China87</td>
<td>Obesity</td>
<td>BMI &gt; 24 kg/m²</td>
<td>243</td>
<td>69.3</td>
<td>12</td>
<td>BC</td>
</tr>
</tbody>
</table>

Abbreviations: BC, Body composition and anthropometry; BM, Biomarkers; BMI, Body Mass Index; PF, Physical Function; PH, Physiological; PS, Psychological; QoL, Quality of life; R, Risk for adverse health event and medical conditions; USA, United States of America; WC, Waist circumference.
Sample size refers to all randomized participants.
*aTotal sample size was 1176 but only 742 participants had BMI > 30 kg/m².
*bThis study included a subsample of participants with obesity.

of life (n = 13, 24.1%), psychological (n = 10, 18.5%), risk for adverse health event and medical conditions (n = 5, 9.3%), medications (n = 3, 5.6%), pain (n = 4, 7.4%), and sleep (n = 2, 3.7%).

The bubble chart (Figure 2) shows in all outcome domains a higher number of studies and assessed outcomes for the obesity compared to the SO phenotype. In four outcome domains, no studies on the SO phenotype were available. In all outcome domains, the number of studies and the number of assessed outcomes declined with advancing time of measurement. The domain with the highest number of reported outcomes was the biomarkers domain, and the domains with the lowest numbers were medications and sleep (Table 3).

3.2.1 Physical function

In total, 42 different outcomes (Table 3) were reported in the domain physical function of which 19 (45.2%) were reported only once (Table S5). Outcome type was mostly performance outcome (PerFO) which was categorized into lower extremity functional performance, performance-based (Instrumental) Activities of Daily Living (IADLs), mobility, balance (static and dynamic), flexibility, strength (functional, power, and maximal), aerobic capacity, endurance, and fine motor skills. Few studies administered Patient-Reported Outcomes (PROs) that were categorized into IADLs, lower extremity functional performance, and osteoarthritis-specific physical function. Figure S1 and Table S5. The three most frequently reported outcomes in obesity RCTs were gait speed (n = 16), VO2max/peak (n = 16), and the Short Physical Performance Battery (SPPB) (n = 11) (Figure S2). The most frequently reported outcomes in SO RCTs were grip strength (n = 6), gait speed (n = 6), and chair rise (n = 4) (Figure S3). None of the SO RCTs measured VO2max/peak, and only one trial measured endurance by the 2-min step test (Figure S4). Outcomes were most frequently reported between 12 and 26 weeks.

Applied methods and given units differed for 16 of the 42 outcomes. For instance, the SPPB was assessed as SPPB, modified SPPB, and expanded SPPB; and different score ranges were reported: 0–12, 1–12, and 0–4.

3.2.2 Body composition and anthropometry

The included RCTs reported 120 outcomes related to body composition and anthropometry (Table 3). Of these, 85 (70.8%) were reported...
only once (Table S6). The domain was categorized into body mass, fat, muscle, bone, cardiovascular, and hepatic outcomes (Figure S4 and Table S6). All reported outcomes were clinician-reported outcomes (CRO). Most frequently reported outcomes in the obesity RCTs were body weight (kg) \((n = 40)\), fat-free mass (kg) \((n = 31)\), and fat mass (kg) \((n = 27)\) (Figure S5), usually reported between 12 and 19 weeks of intervention. Fat mass (%) \((n = 6)\), fat-free mass (kg) \((n = 6)\), and appendicular lean mass (kg) \((n = 5)\) were most frequently reported in RCTs in SO (Figure S6), reported between 20 and 26 weeks of intervention. Either body weight or BMI was reported in 44 RCTs (81.5%). Body weight was not reported in any of the included SO RCTs; however, BMI was reported in one trial. Bone mineral density and content were assessed in six obesity RCTs at times between 12 and 52 weeks and in one SO RCT at the lumbar spine after 26 weeks.

Several technologies (e.g., dual-energy X-ray absorptiometry, bioelectrical impedance analysis, magnetic resonance imaging, hydrostatic weighing, air displacement plethysmography, or computed tomography scan) were used to assess body composition outcomes.

### 3.2.3 | Biomarkers

The biomarkers domain had the highest number of different outcomes \((n = 190)\) (Table 3), with 134 (70.5%) of them assessed only once (Table S7). The outcome categories were blood lipids, glucose metabolism, inflammation, hormones, vitamins, bone metabolism, kidney, liver metabolism, plasma proteins, proteins of skeletal muscle, and muscular health (Figure S7 and Table S7). The vast majority of the biomarkers were measured in the blood \((n = 143)\). Few were also measured in saliva \((n = 1)\), body tissues \((n = 31)\), and breath \((n = 1)\). The most frequently measured outcomes in obesity RCTs were glucose \((n = 24)\), HDL cholesterol \((n = 22)\), and triglycerides \((n = 21)\) (Figure S8); triglycerides \((n = 4)\), cholesterol (HDL \([n = 4]\), total \([n = 3]\), LDL \([n = 3]\)), and C-reactive protein \((n = 3)\) in the SO RCTs (Figure S9). Renal function was measured by three obesity RCTs and by one SO RCT. Bone-related biomarkers were measured in up to four obesity RCTs, however none in individuals with SO. For the same biomarker outcomes, different units were reported, for example, glucose disposal rate was reported in the following units: mg/min/Insulin, mg kg \(^{-1}\) min \(^{-1}\).

### 3.2.4 | Physiological

Thirty outcomes were assigned to the physiological domain, of which 17 (56.7%) were reported only by one trial. It included the following categories: pulmonary and cardiovascular function/exercise performance, energy metabolism, and aerobic fitness. All the reported outcomes were CROs. Only one SO RCT measured physiological...
outcomes related to cardiovascular function (peak systolic flow velocity, end diastolic flow velocity, and wall share rate) (Figure S10). The remaining outcomes were reported by obesity RCTs only. Blood pressure reported between 12 and 19 as well as 20 and 26 weeks (n = 19) was the most frequently reported outcome, while all other outcomes were reported in one, two, or three RCTs.

3.2.5 | Psychological

The psychological domain summarizes 47 outcomes, which were all assessed in obesity RCTs. Of these outcomes, 31 (66.0%) were unique outcomes (Figure S11).

The psychological domain was subdivided into emotional and neuropsychological outcomes. The emotional category, including only PROs, was further divided into 14 subcategories (depressive symptoms, mood, affect, fear of movement, beliefs about how physical activity and work affect and are related to chronic low back pain, feelings related to pain, stress, self-efficacy, loneliness, self-reported psychosocial aspects, perceived benefits and barriers, social support, self-efficacy, and sedentary habits regarding exercise), and mental health with 18 outcomes. The Geriatric Depression Scale (n = 2), the Center for Epidemiological Studies - Depression (n = 2), the Pain Catastrophizing Scale (n = 2), and the Perceived self-efficacy scale (n = 2) were reported in more than one study.

The neuropsychological category, including PerfOs, PROs, and Observer-Reported Outcomes (OROs), was divided into seven subcategories (global, executive function, attention/psychomotor speed, memory, language, visuconstruction, and subjective cognitive complaints) with 29 outcomes. The most frequently reported outcomes in the neuropsychological category were the Mini-Mental State Exam (n = 3), Trail Making Test A and B (n = 3), and semantic verbal fluency (n = 3). Additional eight outcomes were reported twice (Figure S11).

3.2.6 | Quality of life

The quality of life domain consists of 14 different outcomes (Table 3), all being PROs, with a percentage of unique outcomes of 35.7% (n = 5) (Figure S12). It comprises generic (e.g., SF-36 and EQ-5D) and disease-specific (e.g., Impact of Body Weight on Quality of Life, Minnesota Living with Heart Failure Questionnaire, and Kansas City Cardiomyopathy Questionnaire) health-related quality of life outcomes and the measurement of global cognitive judgments of one’s life satisfaction (Satisfaction with Life Scale). The most commonly administered instrument was the SF-36 (n = 11). The outcomes measured in this domain were reported only in obesity RCTs.

3.2.7 | Pain

Pain was measured exclusively in obesity trials using questionnaires (PROs) in four trials with three (75%) of the outcomes being unique. The Patient-Reported Outcomes Measurement Information System - short form (pain subscale) (n = 2) reported between 12 and 19 weeks was used twice (Figure S13).

3.2.8 | Sleep

The domain sleep comprises only two outcomes which were measured between 12 and 26 weeks of intervention by single obesity trials in the form of questionnaires (PRO) (Table S8).

3.2.9 | Medications

The medications domain was self-reported and was assessed in three obesity RCTs (Figure S14). The number of medications taken by the participants and the change in the number of medications following

<table>
<thead>
<tr>
<th>Domain</th>
<th>All RCTs (n = 54)</th>
<th>RCTs addressing obesity (n = 47)</th>
<th>RCTs addressing sarcopenic obesity (n = 7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical function</td>
<td>42</td>
<td>40</td>
<td>16</td>
</tr>
<tr>
<td>Body composition and anthropometry</td>
<td>120</td>
<td>112</td>
<td>26</td>
</tr>
<tr>
<td>Biomarkers</td>
<td>190</td>
<td>183</td>
<td>22</td>
</tr>
<tr>
<td>Physiological</td>
<td>30</td>
<td>30</td>
<td>4</td>
</tr>
<tr>
<td>Psychological</td>
<td>47</td>
<td>47</td>
<td>0</td>
</tr>
<tr>
<td>Quality of life</td>
<td>14</td>
<td>14</td>
<td>0</td>
</tr>
<tr>
<td>Pain</td>
<td>4</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Sleep</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Medications</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Risk for adverse health event and medical conditions</td>
<td>5</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Composite measures</td>
<td>7</td>
<td>7</td>
<td>1</td>
</tr>
</tbody>
</table>

Note: The domains and the number of outcomes reported by the included RCTs are shown. Some outcomes were reported in both phenotype groups.
lifestyle interventions were reported, three times between 12 and 19 weeks and once between 20 and 26 weeks of intervention.

3.2.10 | Risk for adverse health events and medical conditions

Overall six RCTs (12.2%) (four obesity and two SO RCTs) reported metabolic syndrome risk (n = 3), sarcopenia (n = 2), major mobility disability, frailty, and falls (all n = 1) (Figure S15). However, with the exception of one trial for metabolic syndrome risk, sarcopenia and metabolic syndrome risk were reported as z-transformed continuous variables and not as binary outcomes. The two studies reporting the sarcopenia z scores applied different sarcopenia operationalizations (The European Working Group on Sarcopenia in Older People, The Foundation of the National Institutes of Health,88,89).

Seven outcomes were composite measures and could not be categorized into one domain. The “Healthy Aging Index” reported by Shaver et al. (2018) (reported at baseline and after 6 months of intervention) comprises biomarkers plus a cognitive function measure.90 The cardiometabolic risk factor z score reported by Brennan et al. (2020) comprised anthropometric, body composition, and biomarkers outcomes.40 Muscle quality comprised functional and body composition outcomes. The cardiometabolic risk scores (Framingham Risk Score, National Cholesterol Education Program Adult Treatment Panel, International Diabetes Federation Score, and Cardiometabolic Disease Staging Score) were reported by Bragg et al. (2022) (Figure S16).91

The most frequently reported outcomes overall are listed in Table 4.

4 | DISCUSSION

To the best of our knowledge, this is the first evidence map providing an overview of the outcomes and related methodology reported in lifestyle intervention RCTs in community-dwelling older adults with obesity and SO.

We identified 464 different health-related outcomes in the 54 included RCTs relating to 10 domains with physical function, body composition and anthropometry, and biomarkers domains providing the highest number of outcomes.

Maintenance and improvement of everyday functioning are major goals in geriatrics and should be focused on the management of obesity in older people as obesity and SO increase the risk of functional decline and nursing home admissions.10,92,93 It has been demonstrated that older people with obesity achieve poorer scores on physical performance tests.94–96 Forty-two of the identified RCTs measured at least one outcome in the physical function domain, however, using 42 different outcomes. Although there are well-established tests to assess physical functioning in older people, these were not routinely used in included RCTs, especially in those of the obesity phenotype.97 In addition, very few studies reported to use self-reported measures of physical functioning. Studies in individuals with SO focused more on strength and physical performance (especially gait speed), factors that are determinants and diagnostic criteria of this syndrome.12 However, none of the trials investigated the reversibility of the syndrome. Two studies in SO reported a modified, z-transformed sarcopenia index and found improvements in skeletal muscle mass index, grip strength, gait speed, and a decrease in percentage body fat.99,100 The peak oxygen uptake (VO2\text{max}/peak), a measure of aerobic capacity, is an important indicator of cardiorespiratory fitness and a good proxy of health and health decline.98 VO2\text{max}/peak was assessed in 16 RCTs, exclusively in individuals with obesity. While this parameter is resource intense, there are valid

<table>
<thead>
<tr>
<th>Outcome domain</th>
<th>Most measured outcomes in all included RCTs</th>
<th>Number of RCTs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical function</td>
<td>Gait speed</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>VO2\text{max}/peak</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Chair rise</td>
<td>14</td>
</tr>
<tr>
<td>Body composition and anthropometry</td>
<td>Body weight (kg)</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>Fat-free mass (kg)</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>Fat mass (kg)</td>
<td>30</td>
</tr>
<tr>
<td>Biomarkers</td>
<td>HDL cholesterol</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>Glucose</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>Triglycerides</td>
<td>25</td>
</tr>
<tr>
<td>Physiological</td>
<td>Blood pressure</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>Forced vital capacity</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Forced expiratory volume in 1 s</td>
<td>4</td>
</tr>
<tr>
<td>Psychological</td>
<td>Mini-Mental State Examination</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Trial Making Test A and B</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Semantic verbal fluency</td>
<td>3</td>
</tr>
<tr>
<td>Quality of life</td>
<td>SF-36 Short Form (physical component score)</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>SF-36 Short Form (mental component score)</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>SF-36 Short Form (vitality subscale)</td>
<td>5</td>
</tr>
<tr>
<td>Sleep</td>
<td>Pittsburgh Sleep Quality Index</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Patient-Reported Outcomes Measurement Information System (PROMIS)-short form (Sleep)</td>
<td>1</td>
</tr>
<tr>
<td>Pain</td>
<td>Patient-Reported Outcomes Measurement Information System (PROMIS)-short form (Pain)</td>
<td>2</td>
</tr>
<tr>
<td>Medications</td>
<td>Medication change*</td>
<td>3</td>
</tr>
<tr>
<td>Risk for adverse health event and medical conditions</td>
<td>Metabolic syndrome risk</td>
<td>3</td>
</tr>
</tbody>
</table>

Note: The most frequently reported outcomes over all groups and times are reported.

*Medication change is the only outcome in this domain.
Consequently, changes in fat mass and although changes in bone markers are detectable early, the number of medications taken is high in older people. Therefore, it only few trials and... Pain and sleep are also determinants of quality of life, and both are associated with obesity and aging. In addition, pain is a barrier to performing everyday activities and losing body weight. As these domains are directly linked to obesity and phenotype. They constitute mostly PROs and are recommended for use in RCTs. As these domains are directly linked to obesity and aging, this is an important research gap that needs to be addressed in future studies.

Obesity could affect the patients' mental health leading, among others, to depression, stress, and low self-esteem. Therefore, it is not surprising that obesity is associated with lower quality of life. The body weight stigma not only affects the patient's mental health but is also impacting their social participation. Social participation was not an outcome in any included RCT. Considering that advanced age is also linked to loneliness, older people with obesity are at particular risk. Pain and sleep are also determinants of quality of life, and both are associated with obesity and aging. In addition, pain is a barrier to performing everyday activities and losing body weight. The number of medications taken is high in older people with obesity. In diabetic adults with overweight and obesity, it has been shown that body weight loss is associated with a reduction in antidiabetic and antihypertensive drugs.

Heterogeneity was not only introduced by using diverse outcomes but also by differences in applied methodology within outcomes. In addition, different units were reported for similar tests, which were not always convertible. For instance, gait speed (habitual and fast/from standing and flying start) was measured over distances from 4 to 400 m making their comparison questionable and underlining the necessity for harmonization. In addition, different units were reported for a similar test, which were not always convertible. Moreover, statistical reporting of outcomes differed. Some RCTs reported the outcomes as post-intervention values, while others calculated change values. Without doing imprecise assumptions, this causes the inclusion of fewer trials in meta-analysis, leading to data loss. In a few instances, no data were presented for non-significant findings. Establishing recommended outcome methodologies specific for older adults with obesity is needed and would improve the quality of pooled analyses and subsequent recommendations for clinical practice and research.

The timing of measurement varied between RCTs; however, in the majority of the studies, outcomes were assessed between 12 and 26 weeks after randomization. Since only four RCTs provided data beyond 52 weeks, almost no information on the sustainability of lifestyle interventions in older people exist. Only few trials and none in individuals with SO included follow-up assessments after the respective interventions. Specific outcomes, such as bone structural outcomes, keep changing even after 12 months of body weight reduction, and bone loss is an adverse reaction of body weight loss in older adults, although changes in bone markers are detectable earlier. Importantly, lifestyle interventions for body weight loss commonly include behavioural change strategies. Longer interventions, later follow-up times, and weight maintenance studies are required to evaluate the long-term effects and sustainability of outcomes assessed.

As with any study, we acknowledge the limitation of our evaluation. First, due to the diverse operationalizations of SO, people likely differed in their characteristics, thus influencing the number of studies included as well as the number of outcomes and potentially increasing heterogeneity in outcomes. For instance, studies were excluded due to applied cutoffs that could not be validated. Second, we did not extract any behavioural outcomes (healthy food choices, less sedentary behavior, nutrient intake, etc.) as they were not considered direct health outcomes. In most studies, these were applied to monitor study compliance and often not reported as baseline and post-intervention values. However, behavioural changes are crucial for the long-term efficacy of lifestyle interventions and, thus, for the health of individuals.

This work may help in the planning of future RCTs by informing researchers about the outcomes that have been measured and their methodology. The mapping can initiate the closing of revealed research gaps. Further, it provides the base for a consensus process of clinicians, researchers, and patients to select the most relevant outcomes to be included in a COS, considering validity, reliability, responsiveness, feasibility, and cost factors in different settings. The consensus-based “Accumulating Data to Optimally Predict obesity Treatment” project identified a standard set of about 50 core measures or factors that can be analyzed across studies to better understand the variation in response to obesity treatments. While they put a wider focus on obesity-related measures, including...
environmental and behavioural factors, no specifics of (sarcopenic) obesity in older people, such as functional status, disability, and loss of bone mass, were considered. For the selection of outcomes in a COS for obesity, it is important to consider age, frailty, and functional status. The majority of the RCTs (92.6%) included in this evidence map were conducted in “young” older adults (aged 60–74 years), while only 7.4% of the RCTs reported a mean age of 75 years and older. Similarly, only few RCTs made objectively measured or self-reported functional limitations an inclusion criterion.55,59,70,81–83 Regarding outcomes, floor effects should be considered in RCTs in functionally impaired, older individuals. Contrary, measures such as the SPPB may lead to ceiling effects in young-old, functionally intact individuals and are thus not best suited as outcome measures.97 Frail older people with obesity, related comorbidities, and functional impairments could benefit a lot from lifestyle interventions. However, exercise may not be effective caused by reduced trainability in individuals who are vulnerable to changes in homeostasis.120 It may hence be that results from relatively young-old and unimpaired samples cannot be transferred to individuals of advanced age and with functional limitations.

5 | CONCLUSION

This first evidence map on health-related outcomes of lifestyle interventions in older people with obesity or SO displayed a high clinical and methodological heterogeneity regarding used outcomes, their methodology, and reporting. Research gaps include the lack of reporting outcomes over longer periods and the addressing of several domains, such as quality of life and psychosocial outcomes, especially in the SO phenotype. In addition, studies in people aged 70 years and older and in individuals with functional impairments are scarce. Considering the high prevalence and increasing incidence of obesity in older people worldwide, harmonization of outcomes and the development of a COS is highly warranted. This would enable high-quality evidence syntheses to derive evidence-based guidelines and optimize treatment.

ACKNOWLEDGMENTS

This work was funded by the German Federal Ministry of Education and Research (BMBF) grant number (01KG1903). Some preliminary results of this evidence map have been presented in German at the online congress of the German Geriatrics Society (02-04.09.2021) https://doi.org/10.1007/s00391-021-01948-x.pdf. We thank Edith Motschall for peer reviewing the search strategy. The present work was performed in partial fulfillment of the requirements for obtaining the degree Dr. rer. Biol. Hum (Doctoral Degree in Human Biology) for Isabel Galicia Ernst. Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST

IGE, GT, LS, HK, RK, WK, NSB, MV, DVo, EK, DS, and CCS have no conflict of interest to declare. JB has no conflicts of interest related to the submitted manuscript. Outside the submitted work, JB has received support: NIH-K23-AG-051681; grants: K23-AG-051681, R01-AG-067416, R41-AG-071290, R01-AG-058615, U01-AG-071450, and 1UG1HD107692; payment or honoraria for lectures or presentations: Obesity Medicine Association, Wake Forest Baptist Health, Hong Kong Research Grants Council, NIH Reviewer Payment, and Teaching Kitchen Collaborative. Patents planned, issued, or pending: No. 62/672,827. Participation on a Data Safety Monitoring Board or Advisory Board: Nutrition for precision health and other financial or non-financial interest: Equity in SynchroHealth LLC, remote monitoring company.

DVI has no conflicts of interest related to the submitted manuscript. Outside the submitted work, DVI has received grants: RO1-DK109950, CX002161, Biophytis and has received payments from the Clinical Trials Advisory Panel to the NIH-National Institute on Aging and the DSMB for multiple NIA clinical trials for his participation on a data safety monitoring board or advisory board.

ORCID

Isabel Galicia Ernst https://orcid.org/0000-0001-5936-638X
Gabriel Torbahn https://orcid.org/0000-0003-1463-3119
Lukas Schwingshackl https://orcid.org/0000-0003-3407-7594
Helge Knüttel https://orcid.org/0000-0002-2654-6517
Robert Kob https://orcid.org/0000-0002-0406-092X
Wolfgang Kemmler https://orcid.org/0000-0003-3515-0669
John A. Batsis https://orcid.org/0000-0002-0845-4416
Dennis T. Villareal https://orcid.org/0000-0003-1365-7960

REFERENCES


SUPPORTING INFORMATION
Additional supporting information can be found online in the Supporting Information section at the end of this article.