Abstract

Objective: A systematic review was done of the evidence on yoga for improving balance.

Design: Relevant articles and reviews were identified in major databases (PubMed, MEDLINE®, IndMed, Web of Knowledge, EMBASE, EBSCO, Science Direct, and Google Scholar), and their reference lists searched. Key search words were yoga, balance, proprioception, falling, fear of falling, and falls. Included studies were peer-reviewed articles published in English before June 2012, using healthy populations. All yoga styles and study designs were included. Two (2) raters individually rated study quality using the Downs & Black (DB) checklist. Final scores were achieved by consensus. Achievable scores ranged from 0 to 27. Effect size (ES) was calculated where possible.

Results: Fifteen (15) of 152 studies (age range 10–93, n=688) met the inclusion criteria: 5 randomized controlled trials (RCTs), 4 quasi-experimental, 2 cross-sectional, and 4 single-group designs. DB scores ranged from 10 to 24 (RCTs), 14–19 (quasi-experimental), 6–12 (cross-sectional), and 11–20 (single group). Studies varied by yoga style, frequency of practice, and duration. Eleven (11) studies found positive results ($p<0.05$) on at least one balance outcome. ES ranged from -0.765 to 2.71 (for 8 studies) and was not associated with DB score.

Conclusions: Yoga may have a beneficial effect on balance, but variable study design and poor reporting quality obscure the results. Balance as an outcome is underutilized, and more probing measures are needed.

Introduction

One out of three adults age 65 and older fall each year. Older adults who fall are at greater risk for nonfatal injuries, hospitalization, mortality, and decreased independence. Adults may also develop a “fear of falling” (FOF), which may further limit their activity level, increase anxiety, and reduce confidence. The prevalence of FOF among community-dwelling elderly (>60 years) without a history of falls has been reported to be up to 65%. Prevalence can reach up to 92% for those with a history of falls.

Impaired balance and FOF are often the result of multiple shared risk factors, such as psychosocial factors, self-reported health status, and physical fitness. The multifactorial nature of balance deficits and FOF predicates a multifaceted approach. Yoga is a strong candidate for therapeutic intervention, since it provides a comprehensive, integrated approach that can address multiple risk factors at once. Yoga is a mind–body practice stemming from ancient Indian philosophy that has gained popularity in the United States in the past decade. According to the National Center on Complementary and Alternative Medicine, yoga is one of the top 10 complementary health practices used among U.S. adults and has been reported to increase wellbeing and quality of life for the general population.

Yoga is low-impact, can be easily implemented regardless of age or level of experience, and can be modified to suit an individual’s needs. The practice has been shown to have a low rate of side-effects, low risk of injury, and no known interactions with prescription medications.

Adopting yoga as a strategy may aid in the development of awareness of the body through strengthening poses and breathing techniques as it assumes a stable stance or moves in space. Yoga may also boost confidence and reduce anxiety. To investigate these benefits, outcome measures are needed that probe objective and subjective qualities of balance if they are to be useful in future clinical settings. Despite yoga’s popularity, very few studies have been conducted on the purported benefits of yoga for balance. To date, there have been no systematic reviews surveying the quality of available research for yoga and balance.
<table>
<thead>
<tr>
<th>Source</th>
<th>Study design</th>
<th>Study population</th>
<th>Study N (M/F)</th>
<th>Yoga (n)</th>
<th>Control</th>
<th>Mean age (range, mean [SD])</th>
<th>Yoga style</th>
<th>Frequency of practice</th>
<th>Outcome measures</th>
<th>Primary or secondary outcome</th>
<th>Study p-value</th>
<th>Effect size δ</th>
<th>DB score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hart, 2008</td>
<td>RCT</td>
<td>Adults</td>
<td>21 (10/11)</td>
<td>10</td>
<td>Waitlist</td>
<td>Yoga 29 (6), control 26 (7)</td>
<td>Bikram, unspecified style (poses listed)</td>
<td>1.5-h session, 3× a wk for 8 wks</td>
<td>OLS</td>
<td>Secondary</td>
<td>p&lt;0.05b</td>
<td>Data not reported</td>
<td>18</td>
</tr>
<tr>
<td>Hotkar, 2009</td>
<td>RCT</td>
<td>Male students</td>
<td>46 (All M)</td>
<td>23</td>
<td>Non-waitlisted control</td>
<td>11–14 yrs; Yoga 71.5 (4.9), exercise 73.6 (5.1), control 71.2 (4.4)</td>
<td>Iyengar (poses not listed)</td>
<td>1.5-h session/1× per wk with daily home practice encouraged, for 6 months</td>
<td>OLS</td>
<td>Secondary</td>
<td>p&lt;0.05</td>
<td>Yoga vs. control: 0.407</td>
<td>12</td>
</tr>
<tr>
<td>Oken, 2006</td>
<td>RCT</td>
<td>Older adults</td>
<td>135 (34/101)</td>
<td>44</td>
<td>2 control groups: Exercise &amp; Waitlist</td>
<td>30–75 yrs; Hatha (poses listed)</td>
<td>5× a wk for 8 wks; minutes not listed</td>
<td>TUG</td>
<td>Secondary</td>
<td>n.s.</td>
<td>SD not reported</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Singh, 2010</td>
<td>RCT</td>
<td>Adults</td>
<td>80 (Gender not reported)</td>
<td>20</td>
<td>3 control groups: 2 Active &amp; No treatment</td>
<td>30–75 yrs; 59.23 (12.94)</td>
<td>Wooden rail, Bent on Learning Protocol (poses not listed)</td>
<td>5× a wk for 12 wks</td>
<td>OLS</td>
<td>Secondary</td>
<td>p&lt;0.05</td>
<td>0.381</td>
<td>14</td>
</tr>
<tr>
<td>Van Puymbroeck, 2007</td>
<td>RCT</td>
<td>Adults</td>
<td>17 (4/9)</td>
<td>9</td>
<td>Non-waitlisted control</td>
<td>10–13 yrs; Hatha (poses not listed)</td>
<td>45 min session, 3× a wk for 3 yrs</td>
<td>Wooden rail, Bent on Learning Protocol (poses not listed)</td>
<td>5× a wk for 12 wks</td>
<td>OLS</td>
<td>Secondary</td>
<td>p&lt;0.05</td>
<td>0.613</td>
</tr>
<tr>
<td>Bera, 1998</td>
<td>Quasi-experimental</td>
<td>School boys</td>
<td>88 (All M)</td>
<td>50</td>
<td>Non-waitlisted control</td>
<td>10–13 yrs; Yoga 10.4 (0.8), control 10.2 (0.6)</td>
<td>Bent on Learning Protocol (poses not listed)</td>
<td>5× a wk for 12 wks</td>
<td>OLS</td>
<td>Secondary</td>
<td>p&lt;0.05</td>
<td>0.765</td>
<td>15</td>
</tr>
<tr>
<td>Berger, 2009</td>
<td>Quasi-experimental</td>
<td>4th and 5th grade students</td>
<td>71 (24/47)</td>
<td>39</td>
<td>Active</td>
<td>Yoga 10.4 (0.8), control 10.2 (0.6)</td>
<td>Bent on Learning Protocol (poses not listed)</td>
<td>5× a wk for 12 wks</td>
<td>OLS</td>
<td>Secondary</td>
<td>p&lt;0.05</td>
<td>0.64</td>
<td>19</td>
</tr>
<tr>
<td>Silver, 2005</td>
<td>Quasi-experimental</td>
<td>Adults</td>
<td>30 (4/9)</td>
<td>6</td>
<td>Non-waitlisted control</td>
<td>28.7 (6.34)</td>
<td>Bent on Learning Protocol (poses not listed)</td>
<td>5× a wk for 12 wks</td>
<td>OLS</td>
<td>Secondary</td>
<td>p&lt;0.05</td>
<td>0.143</td>
<td>15</td>
</tr>
<tr>
<td>Zettergren, 2011</td>
<td>Quasi-experimental</td>
<td>Older female adults</td>
<td>16 (3/13)</td>
<td>8</td>
<td>Non-waitlisted control</td>
<td>79–93 yrs; Yoga 84 (4.69), control 81.25 (4.94)</td>
<td>Bent on Learning Protocol (poses not listed)</td>
<td>5× a wk for 12 wks</td>
<td>OLS</td>
<td>Secondary</td>
<td>p&lt;0.05</td>
<td>0.066</td>
<td>15</td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Source</th>
<th>Study design</th>
<th>Study population</th>
<th>Study N (M/F)</th>
<th>Yoga (n)</th>
<th>Control</th>
<th>Mean age (range, mean [SD])</th>
<th>Yoga style</th>
<th>Frequency of practice</th>
<th>Outcome measures</th>
<th>Primary or secondary outcome</th>
<th>Study p-value</th>
<th>Effect size da</th>
<th>DB score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hakim, 2010</td>
<td>Cross-sectional</td>
<td>Older adults</td>
<td>52 (7/45)</td>
<td>11</td>
<td>2 control groups: t’ai chi &amp; no treatment</td>
<td>Yoga 73.1, t’ai chi 74.1, control 76.7</td>
<td>Unspecified</td>
<td>Enrolled at least 8 wks, Duration and frequency not reported.</td>
<td>FAB</td>
<td>Co-primary</td>
<td>p = 0.001</td>
<td>Yoga vs. control: 1.216</td>
<td>12</td>
</tr>
<tr>
<td>Sierpowska, 2006</td>
<td>Cross-sectional</td>
<td>Older female adults</td>
<td>20 (All F)</td>
<td>11</td>
<td>Active</td>
<td>60–74 yrs</td>
<td>Unspecified</td>
<td>1-h session, 2× wk, (duration not specified)</td>
<td>TUG</td>
<td>Primary</td>
<td>n.s.</td>
<td>Yoga vs. control: 0.475</td>
<td>6</td>
</tr>
<tr>
<td>dos Santos Alves, 2006</td>
<td>One group, pre–post</td>
<td>Older female adults</td>
<td>30 (All F)</td>
<td>30</td>
<td>N/A</td>
<td>59–70 yrs; 64.52 (3.78)</td>
<td>Hatha</td>
<td>1-h session, 3× wk for 12 wks</td>
<td>Williams &amp; Greene Test of Physical-Motor Function</td>
<td>Secondary</td>
<td>p &lt; 0.05</td>
<td>N/A</td>
<td>11</td>
</tr>
<tr>
<td>Chen, 2008</td>
<td>One group, pre–post</td>
<td>Female senior adults</td>
<td>16 (All F)</td>
<td>16</td>
<td>N/A</td>
<td>60–86 yrs; 68.93 (9.14)</td>
<td>Silver Yoga</td>
<td>70-min session, 3× a wk for 4 wks</td>
<td>TUG</td>
<td>Secondary</td>
<td>p &lt; 0.05</td>
<td>N/A</td>
<td>19</td>
</tr>
<tr>
<td>Hewett, 2011</td>
<td>One group, pre–post</td>
<td>Adults</td>
<td>51 (10/41)</td>
<td>51</td>
<td>N/A</td>
<td>20–54 yrs; 31.57 (9.29)</td>
<td>Bikram</td>
<td>Minimum of 80% attendance for 8 wks with a goal of 3 a wk</td>
<td>OLS</td>
<td>Secondary</td>
<td>p &lt; 0.01</td>
<td>N/A</td>
<td>16</td>
</tr>
<tr>
<td>Schmid, 2010</td>
<td>One group, pre–post</td>
<td>Older adults</td>
<td>14 (Gender not reported)</td>
<td>14</td>
<td>N/A</td>
<td>78.36 (8.75)</td>
<td>Hatha (poses listed)</td>
<td>75-min session, 2× wk for 12 wks</td>
<td>Illinois Fear of Falling Measure</td>
<td>Co-primary</td>
<td>n.s.</td>
<td>N/A</td>
<td>20</td>
</tr>
</tbody>
</table>

Note: For consistency across studies, the sign (+/-) of the effect-size (ES) for Silver (2005) was reversed since lower scores meant better outcomes (reverse scoring).

*ES calculation was based on between-group post-score comparisons where possible.

The p-value reported is for the pre–post change in the yoga group (i.e., not between group tests).

The study included a control group that was not waitlisted and therefore did not receive yoga.

SD, standard deviation; DB, Downs and Black quality rating checklist; RCT, randomized controlled trial; OLS, One-Leg Stand; TUG, Timed Up and Go; A/P, anterior to posterior; M/L, medial to lateral; BBS, Berg Balance Scale; ABC, Activities-specific Balance Confidence Scale; FAB, Fullerton Advanced Balance Scale; N/A, not applicable; n.s., not significant.
Measuring falls can be very difficult and resource intensive, as it requires prospective monitoring, usually for at least a year. In a research setting, clinical and biomechanical measures of balance are often used as a predictive surrogate for falls.\(^\text{21-23}\) Balance can be categorized into static and dynamic balance. Static balance is the ability to maintain the center of gravity within a base of support in a quiet upright position during standing or sitting. Dynamic balance involves maintaining an upright posture while the center of gravity is moving outside the base of support (for example, in walking).\(^\text{22}\) Several screening and assessment tools exist in both clinical practice and research settings that vary in cost and ease of implementation; for a more comprehensive review see Yim-Chiplis and Talbot, Duncan and Studenski, and Whitney et al.\(^{22-24}\) Static measures of balance may include the One-Leg Stand (OLS)\(^\text{25}\) or the Romberg test (quiet standing with both feet together).\(^\text{22}\) The Timed-Up-and-Go (TUG)\(^\text{26}\) is an example of dynamic balance wherein a participant is asked to stand up from a chair, walk a short distance, turn around, return and sit down. The Berg Balance Scale (BBS)\(^\text{27}\) and the Fullerton Advanced Balance Scale (FAB)\(^\text{28}\) evaluate both static and dynamic activities to derive a composite score. FOF is commonly operationalized as “low perceived self-efficacy” in one’s ability to avoid a fall during routine activities of daily living.\(^\text{29,30}\) Subjective measures for FOF and perceived self-efficacy (or confidence) include the Activities-specific Balance Confidence scale (ABC)\(^\text{31}\) and the Illinois Fear of Falling scale,\(^\text{30}\) among others. Many of these measures are scored on a timed or rating scale offering, an aggregated glimpse of balance function as a whole. However, maintaining equilibrium is a consequence of several sensory systems working in concert: the visual, somatosensory (i.e., proprioception), and vestibular systems.\(^\text{32}\) The modified clinical test of sensory interaction on balance evaluates how well the participant is using sensory inputs—somatosensory (firm or unstable surface) and visual (eyes open or closed)—when one or both systems are compromised.\(^\text{32}\) Instability in any condition suggests a sensory problem. Standing in a quiet, static position on a force platform, one can measure postural stability as a function of ground-reaction forces.\(^\text{33}\) Using balance tests that challenge sensory inputs may provide more sensitive measures of balance that are more informative. The determinant of which balance outcome to use in a study should depend on the population of interest and the research question.

The goal of this review is to survey the literature for studies evaluating the benefits of yoga for balance and FOF in a healthy population and to summarize the outcome measures used to determine such benefit. The focus is on a healthy population to first gauge the evidence for safety and efficacy across all age ranges. Secondly, the outcome measures used are summarized in order to determine a baseline reference for balance in healthy populations. Although frail, elderly adults are at greater risk for falls than community-dwelling, older adults,\(^\text{3}\) it is important to assess the potential for balance improvement through yoga as a preventive measure. Indeed, falls have been reported to occur in at least one quarter of the community-dwelling, elderly population.\(^\text{3}\) Clinical populations are not included in this review because the source of balance impairment and its treatment can vary widely, from stroke to Parkinson’s to vestibular disorders.

### Methods

#### Data sources

A comprehensive review of major electronic databases was conducted for articles published before June 2012. The computerized bibliographic databases searched included PubMed, MEDLINE,\(^\text{36}\) IndMed, Web of Knowledge, EMBASE, EBSCO, Science Direct, and Google Scholar. Reference lists from relevant articles and published reviews were manually searched for additional studies. The key search words included yoga paired with an outcome measure that included balance, proprioception, falling, fear of falling, or falls.

#### Inclusion and exclusion criteria

Peers-reviewed articles published in English measuring static or dynamic balance outcomes, balance confidence or self-efficacy, or “fear of falling” in a healthy population due to a yoga intervention were included in the review. All yoga styles were included. Although seated meditation and yogic breathing are key components of yoga, only styles that also incorporate physical yoga postures were included. Randomized (RCTs) and nonrandomized (quasi-experimental, cross-sectional and single group pre-post designs) studies were included in order to broaden the sample of yoga studies with the selected outcomes.

#### Data extraction and quality assessment

Two raters (authors PEJ and AFN) independently evaluated the selected studies using a modified Downs and Black (DB) quality-rating checklist.\(^\text{34}\) Final scores were achieved by consensus. The DB used in this review is considered an appropriate tool\(^\text{35}\) for the purpose of evaluating RCTs and nonrandomized studies that include both internal and external validity (Table 1). Total achievable scores ranged from 0 to 27. According to Roland,\(^\text{36}\) a score $\geq 16$ is considered high and a score $\leq 10$ represents a low score, with moderate scores in between. The final item on the DB checklist, relating to power calculations, was omitted due to lack of clarity in scoring despite contacting the author (Downs) and consulting a biostatistician for clarification.\(^\text{36,37}\)

Instead, the effect size (ES) was calculated as a tool to determine whether a given study had the power to establish clinically important effects; this is also a condition for comparisons across studies. ES is based on differences between two groups; therefore, single-group pre–post designs were not included in this analysis. ES is insensitive to sample size and provides an estimate of the magnitude of the differences observed between groups.\(^\text{38}\) Cohen’s $d$ was chosen to be used as a measure of the total treatment effect size (ES); this quantity is defined as the difference between mean post-scores for each group divided by the pooled standard deviation.\(^\text{38}\) By using the total treatment effect, the ES is not influenced by potential confounders in change scores (e.g., regression to the mean, which can result in an overestimation of the results).\(^\text{39}\) Per criteria in Ferguson (2009),\(^\text{38}\) a score of 0.41 corresponds to a minimally
significant effect, 1.15 corresponds to a moderate effect, and 2.70 or greater corresponds to a strong effect. The ES was calculated for each balance outcome reported and regardless of whether the study reported its outcomes as being significant. A meta-analysis was not conducted due to the limited number of available studies with similar outcome measures and because of overall poor methodological quality.

Results

Search flow

The database search yielded 152 studies published before June 2012 (Fig. 1). Duplicates, studies with clinical populations, and studies without objective or subjective measures of balance as defined in the methods were excluded \((n=117)\). Abstracts, dissertations, and letters to the editor were also excluded \((n=16)\). The remaining articles were reviewed and four additional studies were removed. Chen et al.\(^40\) measured static balance but did not report the data due to a non-significant result. In addition, Chen reported in the results that more than half of their population had unspecified chronic illnesses. Gauchard et al.\(^41\) and Donahoe-Fillmore et al.\(^42\) used interventions combining yoga and another physical activity (e.g., walking) and therefore, it was not possible to discern from the results the effect of a yoga treatment alone. Hill et al.\(^43\) lumped the results of several experimental groups together, even though the groups were separated by type of treatment. Thus, 15 studies were included for full review: 5 RCTs, 4 quasi-experimental designs, 2 cross-sectional studies, and 4 pre–post designs with 1 group (summarized in Table 1).

Types of participants

The studies included in the review focused on a healthy population, ranging from schoolboys to elder adults. The overall mean age was 53.25 years, with ages ranging from 10 to 93 years old across studies. Eight (8) studies involved an older population of 59+ years.\(^{46,44–50}\) Three (3) studies evaluated balance in adolescents and pre-adolescents.\(^{51–53}\) The sample sizes ranged from 147 to 1356 participants. Total number of participants overall was 687 (with totals of 230 males and 342 females reported). The number of participants that were part of the yoga intervention within a study ranged from 6\(^{54}\) to 51\(^{55}\) (total \(n=343\)). The baseline characteristics of intervention and control groups did not differ significantly for most studies, with the exception of 4 studies that did not present enough information to evaluate group differences at baseline.\(^{58,53,54,56}\)

Yoga style, frequency, and duration

Studies varied by yoga style, frequency of practice, and duration (Table 1). Five (5) studies described using Hatha yoga as the intervention. Hatha yoga is an umbrella term for many of the styles of yoga practiced in the United States. Derivations of Hatha yoga styles were used in five studies: Bikram (2), Kripalu (1), Iyengar (1), and Silver yoga (1). Five (5) studies did not specify the style of yoga, although 2 minimally reported the list of postures used. The sequence, type, and number of postures varied considerably between studies. Five (5) studies minimally included at least one balancing posture (e.g., tree or eagle pose) in the yoga intervention. The duration of the yoga interventions ranged from 4 weeks\(^{57}\) to 3 years.\(^{51}\) Participants practiced 2 to 6 times a week for class times ranging from 45 to 90 minutes. Only 2 studies monitored or encouraged home practice.

Types of outcome measures

A total of 25 balance outcomes were reported from the 15 studies included in the review (Tables 1 and 2). Outcomes evaluating static balance were used in 11 studies, while 6 studies evaluated dynamic balance. Three (3) studies used balance assessments that are a composite of static and dynamic balance abilities (with 1 study reporting the global, static, and dynamic subscale scores). Three (3) studies used

\[\text{FIG. 1.} \text{ Search flow diagram.}\]
<table>
<thead>
<tr>
<th>Balance outcome</th>
<th>Measurement</th>
<th>Cited by</th>
<th>Description</th>
<th>Scoring</th>
</tr>
</thead>
<tbody>
<tr>
<td>Static balance</td>
<td>Timed One-Leg Stand (OLS)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Hart, 2008; Oken, 2006; Berger, 2009; Silver, 2005; Hakim, 2010; Chen, 2008; &amp; Hewett, 2011</td>
<td>Subject stands on preferred leg with eyes open and/or eyes closed</td>
<td>Timed (s)</td>
</tr>
<tr>
<td></td>
<td>Balance Rail&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Hotkar, 2009</td>
<td>No description provided</td>
<td>Timed (s)</td>
</tr>
<tr>
<td></td>
<td>Wooden Rail Test&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Bera, 1998</td>
<td>Subjects stood on a rail, with eyes open, on the preferred foot with the other foot off the floor</td>
<td>Timed (s)</td>
</tr>
<tr>
<td></td>
<td>Bass’ Stick Test&lt;sup&gt;d&lt;/sup&gt;</td>
<td>Singh, 2010</td>
<td>The subject is asked to place the ball of the foot crosswise on a stick (1” × 1” × 12”). The opposite foot is lifted from the floor.</td>
<td>Timed (s)</td>
</tr>
<tr>
<td></td>
<td>Williams &amp; Greene Test of Physical-Motor Function&lt;sup&gt;e&lt;/sup&gt;</td>
<td>dos Santos Alves, 2006</td>
<td>No description provided</td>
<td>Timed (s)</td>
</tr>
<tr>
<td>Dynamic balance</td>
<td>Timed-Up-and-Go (TUG)&lt;sup&gt;f&lt;/sup&gt;</td>
<td>dos Santos Alves, 2006; Zettergren, 2011; Sierpowska, 2006</td>
<td>Subjects stand up from a chair, walk a short distance (~ 8 ft), turn around, return, and sit down</td>
<td>Timed (s)</td>
</tr>
<tr>
<td></td>
<td>Dynamic Leg Swing&lt;sup&gt;g&lt;/sup&gt;</td>
<td>Silver, 2005</td>
<td>Subjects were asked to rhythmically swing the leg in the medial-to-lateral plane (M/L swing) and then again swing the leg in the anterior-to-posterior plane (A/P swing)</td>
<td>Center of pressure</td>
</tr>
<tr>
<td>Composite balance</td>
<td>Berg Balance Scale (BBS)&lt;sup&gt;h&lt;/sup&gt;</td>
<td>Schmid, 2010; Zettergren, 2011</td>
<td>Performance-based scale with 14 items that examine functional movements</td>
<td>5-point ordinal scale. Includes a Global score and dynamic and static subscales.</td>
</tr>
<tr>
<td></td>
<td>Fullerton Advanced Balance Scale (FAB)&lt;sup&gt;i&lt;/sup&gt;</td>
<td>Hakim, 2010</td>
<td>Physical performance-based scale with 10 items that examine static, dynamic, and reactive balance</td>
<td>5-point ordinal scale</td>
</tr>
<tr>
<td>Perceived self-efficacy or fear of falling</td>
<td>Activities-Specific Balance Confidence Scale (ABC)&lt;sup&gt;j&lt;/sup&gt;</td>
<td>Hakim, 2010; Zettergren, 2011</td>
<td>16-item questionnaire in which level of confidence in performing situation-specific activities is rated</td>
<td>Rating (0–100%)</td>
</tr>
<tr>
<td></td>
<td>Illinois Fear of Falling Scale (FOF)&lt;sup&gt;k&lt;/sup&gt;</td>
<td>Schmid, 2010</td>
<td>19-item rating scale was used to assess the level of fear of falling for situation-specific activities</td>
<td>4-point ordinal scale transformed to logit scale</td>
</tr>
</tbody>
</table>

<sup>b</sup>No source provided.
subjective measures of perceived self-efficacy for one’s ability to avoid a fall during routine activities of daily living (e.g., ABC or FOF). Four (4) studies used multiple balance assessments (i.e., objective and subjective measures), and 7 studies used balance as a primary outcome or co-primary measure (only balance measures are reported here). Many of the outcomes are well known and accepted in the research community at large for their accessibility, validity, and reliability (e.g., OLS, TUG, BBS, FAB, ABC, FOF); however, several studies reported balance tests that are less commonly used (e.g., Rail test, Bass’-stick), making it difficult to determine reliability or compare with other outcomes (Table 2). With the exception of Hart and Tracy and Hakim et al., no study included in this review measured the sensory contribution to balance. Hart and Tracy measured OLS with eyes closed reporting ceiling effects, while Hakim et al. used a composite measure (i.e., FAB) of several factors for balance but did not distinguish among the subtasks. Schmid et al. measured falls as an exploratory co-variable and found differences between fallers and non-fallers for the primary outcome measures. Eleven (11) studies found significant results (p < 0.05) on at least one balance outcome (9 static balance, 1 dynamic balance, and 2 composite measures of balance, in Table 1).

Downs and Black quality rating

Downs and Black quality ratings ranged from 10 to 24 (RCTs), 14 to 19 (Quasi-experimental), 6 to 12 (Cross-sectional), and 11 to 20 (pre–post), indicating a diverse range of scores regardless of research design (Table 3). Eight (8) of 15 studies received a high score (≥16), 5 studies achieved moderate scores, and 2 scored poorly with ≤10. A subset of items from the DB detailing internal and external validity is summarized in Table 3. External validity (EV) was generally poor across studies with regard to the selection of study participants. Most studies selected reliable or valid balance measures except where noted. Yoga is ubiquitously available, and therefore all protocols provided evidence of changes in balance through yoga. Fifteen (15) studies of diverse methodology in a healthy population were included. Quality assessment was determined using the DB scale. Additionally, ES was calculated to determine the magnitude of differences. The data from the studies in this review suggest that yoga may have an effect on balance; however, differences in the quality of reporting and methodology limit the interpretability of the results.

Quality of study design

The built-in features of inclusion of a control group and randomization in an RCT are meant to ensure equivalence at baseline, which is required to allow comparison between groups after the intervention. Between-group post-score comparisons (treatment versus control) make it possible to control for natural progression of diseases, spontaneous fluctuation of symptoms, and regression to the mean or placebo effects (i.e., nonspecific effects). The choice for design of an RCT help ensure that the outcomes observed are due to the intervention alone, resulting in higher IV.

Despite the assurance of good study design, there were several violations to IV for the RCTs included in this review. Two (2) studies did not report baseline differences, therefore limiting the interpretation of post-score comparisons between groups. One (1) study did not describe the randomization process and another study did not randomize participants to groups but instead randomized the groups to the treatment or control. Although a potential benefit of the control group is to allow comparison to the treatment group both at baseline and at the end of the study, 2 RCTs only reported pre–post comparisons within groups and did not evaluate differences between groups, introducing potential confounds inherent in evaluating change scores. Two (2) studies included a control group; however, it was not clear whether they were waitlisted. Without waitlisting, a researcher is unable to control for the subjects’ expectations or beliefs about possible benefits that may result from their participation in the study, an effect commonly known as expectation bias. Another threat to IV that challenges the yoga research field is the inability to blind participants to the intervention, since there is no such thing as a “sham” yoga treatment.
<table>
<thead>
<tr>
<th>Source</th>
<th>Study design</th>
<th>DB score</th>
<th>Was the selection of study participants representative of the population of interest?</th>
<th>Was the selection of yoga intervention representative of what happens in real world?</th>
<th>Were the balance outcomes accurate (valid and/or reliable)?</th>
<th>Threats to external validity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hart, 2008</td>
<td>RCT</td>
<td>18</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Consistent, Yes</td>
</tr>
<tr>
<td>Hotkar, 2009</td>
<td>RCT</td>
<td>12</td>
<td>Unable to determine</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Oken, 2006</td>
<td>RCT</td>
<td>24</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Singh, 2010</td>
<td>RCT</td>
<td>10</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Van Puymbroeck, 2007</td>
<td>RCT</td>
<td>20</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Bera, 1998</td>
<td>Quasi-experimental</td>
<td>14</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Berger, 2009</td>
<td>Quasi-experimental</td>
<td>19</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Silver, 2005</td>
<td>Quasi-experimental</td>
<td>15</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Zettergren, 2011</td>
<td>Quasi-experimental</td>
<td>18</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Hakim, 2010</td>
<td>Cross-sectional</td>
<td>12</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Sierpowska, 2006</td>
<td>Cross-sectional</td>
<td>6</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>dos Santos Alves, 2006</td>
<td>One group, pre-post</td>
<td>11</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>N/A</td>
</tr>
<tr>
<td>Chen, 2008</td>
<td>One group, pre-post</td>
<td>19</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Hewett, 2011</td>
<td>One group, pre-post</td>
<td>16</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Schmidt, 2010</td>
<td>One group, pre-post</td>
<td>20</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Only a subset of questions from Downs and Black quality-rating checklist (DB) represented.

*Randomization by drawing slips of paper.

Only results on change scores for each group were reported. Between-group comparisons were not reported.

Outcome measure not adequately described.

Randomization not described.

Homogeneous sample due to exclusion/inclusion criteria therefore difficult to generalize.

Planned modified minimization scheme to maintain balance across multiple stratification variables.

Subjects were divided into groups and groups were randomly allocated to treatment or control. Group assignment procedure unclear.

Stratified randomization.

Large number of dropouts > 30%; large enough dropout that might have made a difference in the analysis.

N/A, not applicable to the design.
Minimally, yoga researchers can blind the assessor; however, only 346,49,50 studies reported doing so. It is possible that due to the preliminary nature of several studies, the authors lacked the resources or personnel to implement these safeguards.

Threats to IV are more pronounced in study designs that are not randomized (i.e., quasi-experimental), where results are taken at a single time point (i.e., cross-sectional designs), and where the design lacks a control group (i.e., single group, pre–post), eliminating the possibility of assessing change between groups. Despite these methodological limitations, results derived from feasibility or pilot studies of non-RCTs designs serve to generate new hypotheses and inform the implementation of larger, more resource-intensive trials. In general, the overall quality of the remaining studies was mixed.

One major criticism of the RCTs included here is their lack of EV.66 EV ensures that the selection of participants is representative of the population, the treatment is generalizable, and the outcomes are clinically relevant. Researchers will generally adhere to inclusion and exclusion criteria, resulting in a more homogeneous population, making it difficult to generalize to other populations with potential comorbidities, for example. The popularization of yoga in the West has made classes readily available and therefore, most yoga study protocols included in this review are representative of what is in use in the general population. Lastly, the selection of outcomes should be reliable and/or valid in order to compare results across studies and to allow generalization to the population at large. In some cases, the reported outcome measures were not commonly used in practice, and thereby are less amenable to review.

Only two of the 15 studies included in the review reported a power calculation to determine sample size,66,47 and no study evaluated effect sizes. Calculating ES can inform the development of future studies, in particular, if the study reported nonsignificant findings. A moderate-to-high ES may be an indicator of a real effect even if the study lacked the power or the sample size to yield significant results. ES was calculated in this review for studies providing sufficient information in the results. Only 2 studies reported significant findings with a corresponding ES in the moderate-to-high range,5,53 yet the moderate quality of study design, as evidenced by their scores on the DB, makes it difficult to draw definitive conclusions for these studies.

**Outcomes**

The most commonly used outcome was the OLS, a measure of static balance that requires a person to stand on one leg. The ability to stand on a single leg is an important predictor of falls: One study demonstrated that the chance of sustaining an injury due to a fall is twice as high if one is unable to perform the OLS for 5 seconds.62 Yoga standing poses that invoke the use of hip abductors and/or adductors, such as tree pose (virksasana), mirror the biomechanical demands required to successfully execute the OLS.83 Therefore, one may hypothesize that practicing poses that challenge static balance would lead to improvements in balance. Ironically, 3 studies that measured balance did not include any standing postures that may lead to improvements, which may have contributed to the reported nonsignificant findings.45,54,56 The selection of yoga style and which poses to include is an important consideration in any study.63–65

Simple balance tests that rely on timing or rating scores are limited in their ability to predict future fall risk.66 It has been suggested that balance tests that include measures of postural stability in the anterior–posterior and mediolateral direction are more predictive of fall risk and may provide outcomes more sensitive to changes after a yoga intervention.66 Silver et al.54 was the only study in this review to include measures of postural stability (i.e., center of pressure); however, they did not report positive results. This may have been due to ceiling effects in a young adult population, or because they purposefully omitted standing balance poses in their intervention, or because of the small sample size (due to a high dropout rate), or because they failed to conduct a baseline comparison.

Measuring dynamic balance, in addition to static balance, can provide another dimension of functional performance; however, common tools such as the TUG may not be sensitive in healthy populations.57 Of the 4 studies that used the TUG in this review, only 2 reported positive results. Whether static yoga postures can improve performance of dynamic timed walking tests has yet to be determined. Furthermore, dynamic tests have shown low predictive ability for falls.22 Yoga has the potential to improve balance for the prevention of falls; however, falls were included as an outcome in only 1 study, and only as an exploratory covariate.47 Another limitation of dynamic balance tests is their inability to detect the source of the balance deficit (i.e., due to visual, somatosensory, or vestibular problems). Yoga is purported to improve somatosensory mechanisms (e.g., proprioception); however, no study in this review evaluated this domain. Two (2) studies45,58 considered the visual contributions to balance by measuring static balance activities with eyes closed. These data were not explicitly reported because the measure was either embedded in a composite score45 or not significant due to ceiling effects.58

Composite measures such as the BBS and FAB include a combination of static and dynamic measures reported as an overall global score on an ordinal scale. One limitation of measures that are scored on an ordinal scale is that they may lack the precision necessary to determine subtle changes in a healthy population (i.e., ceiling effects).67 While an overall score can be useful in the initial assessment of balance impairment, it may also be important to know which aspect of balance is most affected in order to target an intervention. The BBS includes subscales differentiating between static and dynamic performance in its scoring rubric, which can be more informative than the composite score alone. For example, Schmid et al.47 reported the global BBS score as nonsignificant, yet found the static balance subscale changes to be significant \( (p=0.045) \) but not those on the dynamic subscale, suggesting that static balance may be more amenable to a yoga intervention. Zettergen et al.50 reported significance on the global score for BBS but not its subscales, making it difficult to evaluate whether yoga improved static or dynamic balance or both. Similarly, another study45 reported significant effects using the FAB; however, it is unclear which aspects of balance improved.

Objective measures are less sensitive to bias, yet may not capture the experience of the individual, and therefore
relevant subjective measures are important to consider. As an example, adverse effects related to FOF can include depression, self-imposed restriction of daily activities, and a decline in physical fitness. Individuals may experience anxiety in their ability to avoid a fall. It is well recognized that FOF is a risk factor for falls in an aging population. Of the studies conducted in an aging population (>60 years, n=8) only 3 evaluated perceived balance self-efficacy, which is a risk factor for falls in that group. All 5 studies reported nonsignificant results, which may be due to the small sample size in these studies or to the lack of sensitivity in self-report measures. To ensure that patients’ experiences are taken into account, measuring patient-reported outcomes in parallel with objective clinical outcomes will be important for future clinical trials.

Quality of the yoga intervention

Although there is no specific dose response for yoga practice, improvement in measurable outcomes is directly related to the frequency and duration of practice. There was no clear correlation between frequency and duration of practice and significance or ES for studies included in this review. For example, the study by Hotkar (2009), which included a high frequency of practice (6 times/week for 6 weeks), showed significant results and a high ES. In contrast, participants in Singh et al. (2010) practiced 5 days a week for 8 weeks with nonsignificant results. It is possible that the differences in results were due to the differences in outcome measures used or other methodology not reported (i.e., home practice was not part of the intervention nor was adherence reported). The longest study duration (45 minutes 3 times a week for 3 years) found improved balance in a population of young school boys; however, there were limitations in the methodology. Ensuring that the yoga sequence was delivered as specified by the protocol is another issue to consider; however, no study in this review reported oversight on the fidelity of treatment. Several studies that reported positive results did not list the yoga poses used, making it difficult to determine the efficacy of the yoga program for balance in these cases.

Conclusions

Yoga is reputed to improve balance and other psychological factors. This review suggests that yoga may have an overall beneficial effect; however, differences in quality of reporting and study design make it difficult to draw definitive conclusions. Although adverse events were not consistently discussed, none were reported, thus suggesting that yoga may be safe for use in the general population. Considering the number of yoga studies available in the field at large, studies including balance and FOF or perceived self-efficacy as outcome measures in healthy populations are lacking. The number of RCTs was low, and several failed to meet the level of scientific rigor expected in this design. Studies are needed that evaluate specific poses for improving balance, especially those that target balance-affiliated muscles such as hip abductors. The multifaceted nature of balance dictates a need for more sensitive outcome measures that can demonstrate the source of postural instability and the validity of the measurement as a predictor of falls. For example, understanding the somatosensory contribution to balance may lend itself to postulating a mechanism through which yoga can improve balance. Furthermore, more sensitive measures (e.g., posturography) may be less prone to ceiling effects than measures that rely on timing or rating scores (e.g., OLS). Objective measures alone do not reflect the participant’s experience; therefore, subjective outcomes should be included in the research design. In the future, yoga researchers may consider evaluating postures that promote balance and using more sensitive measures in well-designed studies.

Author Disclosure Statement

The authors have no personal or financial interests to disclose.

References


Address correspondence to:
Pamela E. Jeter, PhD
Lions Vision Research Center
Wilmer Eye Institute
Johns Hopkins University
600 North Wolfe Street
Woods 355
Baltimore, MD 21287
E-mail: pamela.jeter@gmail.com