THIRTY YEARS ON: A NARRATIVE REVIEW OF RESEARCH ON STRENGTH TRAINING FOR FEMALE ATHLETES SINCE THE NATIONAL STRENGTH AND CONDITIONING ASSOCIATION’S POSITION PAPER

Authors: Jason Shurley, Ph.D., University of Wisconsin - Whitewater and Jan Todd, Ph.D., University of Texas at Austin

ABSTRACT

In the fall of 1989, the National Strength and Conditioning Association (NSCA) published its position paper on strength training for female athletes. The work ultimately concluded that females responded to strength training programs comparably to male counterparts and should thus be trained in a similar fashion. The authors went on to point out, however, that their conclusion was based on a relatively small body of research. This article reviews the research on strength training related to female athletes since the publication of the NSCA’s position paper. Research since that paper’s publication has largely validated the initial findings with additional data. It also has shown that multiple sets and periodized programs appear to be more effective at improving muscular size, strength, and power. Injury prevention programs, such as those designed to mitigate knee injuries, also have been shown to improve parameters of athletic performance.

INTRODUCTION AND BACKGROUND

In the fall of 1986, the National Strength and Conditioning Association (NSCA) commissioned a special committee to generate a position paper related to strength training for women (Holloway et al., 1989a; Holloway et al., 1989b; McQuilkin & Smith, 1995). The official charge of the “Women’s Position Paper Committee” was to comb the research on resistance training and report whether women should be trained or coached differently than their male counterparts. The committee consisted of six members only one of whom, Meg Ritchie (Stone), was employed full-time as a strength and conditioning coach. Ritchie was joined on the committee by Denise Gaiter who was a graduate student and exercise physiologist at an Arizona fitness resort, Jean Barrett Holloway who taught continuing education classes in strength training at the University of California – Los Angeles, Lori Gilstrap who was also a graduate student and part-time strength and conditioning coach at Georgia Tech, Lynne Stoessel a graduate student in exercise physiology at Auburn University, and Jan Todd a lecturer in Kinesiology at the
University of Texas at Austin and multi-time world and national record holder in powerlifting (Holloway et al., 1989; McQuilkin & Smith, 1995).

By the time the women’s committee was assembled, the NSCA had already published two position papers. Both appeared in the NSCA Journal in 1985 with the first, on use and abuse of anabolic steroids, appearing in April and followed by a paper on prepubescent strength training in August (“National Strength and Conditioning Association: Use and abuse,” 1985; “Official Document: Prepubescent strength training,” 1985). The one-page summary of the organization’s stance on anabolic steroids came out just one year after the American College of Sports Medicine (ACSM) released its own position statement and First Lady Nancy Reagan began her “Just Say No” anti-drug campaign (Seefeldt & Ewing, 1997; “Timeline,” 2007). The statement also came out amidst a torrent of criticism and concern about their use. Sports Illustrated, for example, published a special report on “the steroid explosion” in May of 1985 and described anabolic steroid use as “a spreading wildfire that is touching athletes at every level of sport” (Denham, 1997; Johnson, 1985).

Similarly, the paper on prepubescent strength training followed on the heels of a position statement by the American Academy of Pediatrics (AAP) that discussed weight training and weightlifting in youth and adolescents (American Academy of Pediatrics, 1983). The AAP position separated weight training, which involved machines and pulleys but “no free weights,” from weightlifting which included the competitive lifts performed by both Olympic lifters and powerlifters. Weight training, the pediatricians assured parents, would not lead to muscle-boundness, and would enhance athletic performance. Weightlifting, on the other hand, could result in growth plate fractures. Injuries of the low back, knee, and shoulder were also “common” in individuals who performed these lifts. In defense of their own profession, the NSCA paper stated repeatedly that weight training in even prepubescent subjects was both “safe and efficacious” if properly implemented and supervised (“Official Document: Prepubescent strength training,” 1985).

As with the first two papers, a variety of factors led to the creation of the Women’s Committee and, ultimately, its Position Paper on resistance training in female athletes. The most immediate factor was a dramatic underrepresentation of women in the NSCA. As an example, in 1985 the NSCA sought to create a profile of individuals carrying the association’s Certified Strength and Conditioning Specialist (CSCS) credential; of 140 respondents to that survey, only 4 were women (Baechle, 1985). While women were scarce in the NSCA, they were becoming increasingly prevalent in competitive athletics. Since the passage of Title IX of the Education Amendments of 1972, female participation in athletics at both the high school and collegiate level had spiked dramatically. At the high school level the number of athletes
more than doubled, from 817,000 girls participating in 1972 to 1.8 million in 1986 (National Federation of State High School Associations, 2015). Female collegiate athletes more than tripled from 30,000 to 95,300 during the same span (National Collegiate Athletic Association, 2012; Suggs, 2006).

In addition to traditional high school and collegiate sports, women were also gaining a foothold in strength sports previously reserved only for men. The first competitive powerlifting competition for women in the United States was organized in 1977 with the first national championships in Olympic weightlifting taking place four years later in 1981. It is worth noting that the first women’s world championships in weightlifting did not take place until 1987 and that it was not included on the Olympic program until 2000 (M.H. Stone, Pierce, Sands, & Stone, 2006). In competitive bodybuilding the first women’s championships were held in 1980 and won by Rachel McLish, with Laura Combs being crowned the first Ms. America in the same year (Fair, 2015). Unfortunately for proponents of strength training, it was the image of female bodybuilders which became indelibly etched onto the minds of the American public when it came to the effects of strength training for women.

As discussed in the Position Paper, many women were hesitant to take up serious strength training for fear that it would lead to excessive musculature and create a masculinizing effect (Holloway et al., 1989a). One factor in this belief may have been Cold War imagery of female Eastern-bloc athletes and their frequent portrayals as being of “questionable femininity” and described with such terms as “muscular” and “hefty” (Cahn, 2003; Schultz, 2014; Matelski, 2017). As concrete examples, in the lead-up to the formation of the Women’s Committee, Sports Illustrated’s coverage of East German female swimmers used terms like “strapping” to describe the athletes, consistently referenced their enormity by comparing their heights and weights to American swimmers, and referenced the Communists’ lack of shaving (Kirshembaum, 1976; Neff, 1985; Neff, 1986). Many of those athletes, of course, were later discovered to have been unknowingly doped with anabolic steroids over their years of training (Ungerleider, 2001).

Further, the 1980s saw what journalist Susan Faludi described as a “backlash” against the progress women had made in the prior decade (Faludi, 2006; Schultz, 2014). One manifestation of this, according to historian Jaime Schultz (2014), was the reconstructing of exercise and athletics as “a means to a heterosexay end and not...and end to itself” (p. 126-127). Women of the 1980s were encouraged by mass media to work out and reshape their bodies primarily for aesthetics. It did not matter what a woman could actually do, so long as she looked “fit.” The development of this “to be looked at aesthetic” is described by Schultz as a response to gains made by women on the playing fields and courts of the 1970s and helped foster the aerobic dance craze of the 1980s (Schultz, 2014, p. 131-146).
Aerobic dance programs suited the needs of the fit aesthetic by providing activity sufficient to reduce bodyfat, provided one was a regular practitioner, but not exercise strenuous enough to develop any real muscularity.

Into this milieu of “backlash” against female muscularity, the Women’s Committee began its work reviewing the literature on strength training for female athletes. Major findings of the Position Paper included that: females had similar relative strength to males when body composition was taken into account, women hypertrophy to the same relative degree as males, fiber type is similar between sexes, there was scant evidence demonstrating that normal menstruation hampered athletic performance, and that there was no basis for the concern that narrow female shoulders could not support overhead lifts, though a larger angulation between the femur and tibia (Q-angle) may predispose female athletes to knee injuries (Holloway et al., 1989b). “Due to similar physiological responses,” the authors asserted, “it appears that males and females should train in the same basic way, employing similar methodologies, programs, and types of exercises” (Holloway et al., 1989b, p. 30). Committee members went on to cite the lack of strength training research on female participants as a “major concern” and strongly advocated for both more research and, specifically, research that employed more realistic protocols to the strength training that might actually be employed by strength coaches and female athletes (Holloway et al., 1989b).

The lack of research on the effects of strength training on women generally and female athletes particularly should not have been especially surprising. Research on any type strength training was a niche enterprise until the 1980s, when it became a more legitimate area of interest after the formation of the NSCA (Shurley, J. Todd, & T. Todd, 2017). Moreover, women were often excluded from interventional research in the 1960s and 1970s through both custom and policy. Following birth defects attributed to administration of thalidomide to pregnant women and rare cancers in the daughters of pregnant women who took the hormone diethylstilbestrol (DES), researchers were hesitant to include pregnant or potentially pregnant women in clinical trials (Mastrioanni, Faden, & Federman, 1994). Both Congressional action and Food and Drug Administration (FDA) guidelines by the late 1970s advised against inclusion of pregnant subjects. Eventually this aversion toward female subjects extended to all women of child-bearing age. The result of exclusion of females from clinical research was a significant gap in understanding about a variety of pathologies including osteoporosis, cancer, and cardiovascular disease.

As millions of “Baby Boomers” hit mid-life during this era, women began to clamor for a correction of this knowledge gap (Mastrioanni et al., 1994). In the early 1990s, coinciding with the
publication of the Position Paper, a sea change began to take place regarding the presence of female subjects in research projects. In 1991 the National Institutes of Health (NIH) issued guidance to applicants for their grants that women should be included as subjects and, if they were not, that there was a “clear and compelling” justification for their absence. That same year Bernadine Healy was appointed as the first female head of the NIH and announced a $625 million Women’s Health Initiative that sought to examine the effect of exercise, smoking cessation, dietary and other interventions on osteoporosis, cardiovascular disease, and cancer (Mastroianni et al., 1994). One of the results was an explosion of interventional research on the effects of strength training on osteoporosis. The stage was set, then, to rectify the knowledge gap singled out by the Women’s Committee regarding the effects of strength training on female athletes.

It has been almost thirty years since the publication of the Committee’s work, a milestone which prompts the question, “what has changed?” The number of female athletes participating at National Collegiate Athletic Association (NCAA) institutions across all levels has increased to over 200,000 (National Collegiate Athletic Association, 2016). More than 3.2 million girls participate in high school athletics (National Federation of State High School Associations, 2016). There are approximately 150 women working as strength coaches at the NCAA Division I level, up from 3 in 1991 (J. Todd, Lovett, & T. Todd, 1991; Laskowski & Ebben, 2016). In terms of the practice of strength and conditioning, a recent study indicates that the answer might be that little has changed. Though limited in geographic scope and to high school athletes, a 2012 study found that coaches of male athletes were much more likely than coaches of female athletes (50% to 9%) to require their athletes to lift weights (Reynolds, Ransdell, Lucas, Petlichkoff, & Gao, 2012). Moreover, male athletes were more likely to strength train year round and when female athletes did train were more likely to do “female preferred” activities like Pilates and yoga.

The purpose of this narrative review is to examine research related to strength and conditioning for women since the publication of the NSCA’s position paper in 1989. Articles included in this review were found using SportDiscus, Medline, and PubMed databases. Search terms included combinations of: women, woman, female, strength training, resistance training, and athlete. Several broad areas of investigation emerged. This review will pay particular attention to research related to the strength training which might be relevant for female athletes, as the original Position Paper specifically discussed strength training for female athletes. Criteria for inclusion were the use of resistance training modalities including barbell, dumbbells, pulleys, and strength machines. Further, research was included that focused on variables related to athletic performance such as muscular strength, power, hypertrophy,
aerobic capacity, and throwing velocity. Articles were excluded if they did not separate results and specifically discuss the effects of strength training on female subjects.

Muscular Size, Strength, and Fiber Type

One prominent area of investigation, especially in the 1980s and 1990s, examined how women’s muscles responded to strength training and how that compared to male counterparts. Published in the same year as the position paper, the work of Robert Staron and colleagues (Staron et al., 1989) from the University of Ohio echoed the concern of the Women’s Committee that little research existed on the effect of strength training in women. “Despite the large number of women participating in weight training,” they observed “few studies have thoroughly examined its effect on the development of strength, muscle mass, and fiber composition” (Staron et al., 1989, p. 71) Following 20 weeks of twice-weekly training on squats, leg press, leg extension, and leg curls using 80-85% 1RM for 3 sets, the authors demonstrated significant increases in the cross-sectional area (CSA) of both fast and slow-twitch fibers. Additionally, a shift in fast-twitch fibers away from type IIb toward type IIa was observed following the training (Staron et al., 1989). A follow-up study two years later using some of the same subjects as well as additional untrained subjects demonstrated that detraining for more than 7 months had little effect on CSA but allowed fast-twitch fiber types to shift back toward type IIb (Staron et al., 1991). Maximal strength declined with detraining but did not reach pre-training values. Upon resumption of training, hypertrophy was again observed in both fast-twitch fiber types, along with a shift of those fibers back toward type IIa, and increased strength after as little as 6 weeks of training. Using a similar protocol in study published 3 years later, Staron and colleagues (Staron et al., 1994) observed similar relative increases in strength for both previously untrained male and female subjects. The lack of a significant increase in fiber CSA led the authors to theorize that hypertrophic changes take longer than 6 weeks, though a shift in fast twitch fiber type was observed in as little as 2 weeks. While there appeared to be a difference in hormonal response between male and female subjects, a subject which will be revisited later in this work, the authors concluded that skeletal muscle itself responds similarly to resistance training in men and women.

Consistent with the findings of the Position Paper, Miller and colleagues (Miller, MacDougall, Tarnopolsky, & Sale, 1993) found a similar proportion of fast and slow-twitch fibers and a similar fiber number in both male and female subjects. The male subjects had both higher absolute strength scores and strength relative to lean mass. The authors attributed the higher male strength scores to the higher CSA of male fibers compared to those of female counterparts. Whether the difference in fiber size was due to previous activity or an innate difference was unknown, though the authors speculated that larger
fibers in men may be an innate sex difference. In a study of 20 weeks of strength training for the elbow flexors using untrained subjects, O’Hagan and colleagues (O’Hagan, Sale, MacDougall, & Garner, 1995) found greater relative increases in strength in female subjects. Both male and female subjects had similar increases in muscle CSA. Using a whole-body rather than a single-joint training program, Chilibeck and colleagues (Chilibeck, Calder, Sale, & Weber, 1998) observed increased strength throughout their 20-week program for untrained young women. Of note, however, was that the early strength gains occurred before measurable hypertrophy. Further, hypertrophy was observed earlier in the upper body than the lower body or trunk.

Abe and colleagues (Abe, DeHoyos, Pollock, & Garzarella, 2000) also noted that muscular hypertrophy was likely to occur earlier in the upper body than it was in the lower body for both men and women. Following a 12-week total body resistance training program, untrained male and female subjects made similar relative increases in both muscular strength and hypertrophy. Changes in size and strength followed a similar timeline for both genders. Using resistance training exercises only for the elbow flexors and extensors over the course of 12 weeks, Hubal and colleagues (Hubal et al., 2005) found that men made greater relative gains in muscle size than women. The difference between the sexes was small, 2%, but statistically significant. Women showed marked increases in muscular strength both compared to baseline and relative to men. The authors speculated that this difference in relative strength was attributable to the lower initial strength of the untrained women.

Following reports (Hickson, 1980; Dudley & Djamil, 1985) that endurance training could hamper strength performance several investigators sought to determine whether that was true in female subjects who combined strength and endurance training. Using sedentary young women, Volpe and colleagues (Volpe, Walberg-Rankin, Rodman, and Sebolt, 1993) compared strength gains made by a group of women who only used resistance training to another group that engaged in both strength training and endurance running. Both trained on a combination of Universal and Nautilus weight machines for 9 weeks, performing repetitions ranging from 4-12. The endurance group then ran for 25 minutes following weight training. Neither group significantly altered its weight or body composition though both increased strength similarly. Gravelle and Blessing (2000) combined rowing and strength training using both free weights and Universal machines in their 11-week intervention. Recreationally active women either lifted exclusively, rowed and then lifted, or lifted and then rowed. The strength program in this study used 2-4 sets of 6-10RM on 6 lower body exercises. All groups increased strength significantly and similarly. The primary finding of difference was that improvement in aerobic capacity was smaller when endurance training was performed before strength training, rather than strength
before endurance. That finding, however, may be due to the fact that 1 of the 6 subjects in the group that rowed first experienced a decline in her aerobic capacity, substantially decreasing the group average.

In general, research on female responses to strength training programs bore out the findings of the Position Paper. Women did respond to training similarly to men. Both male and female subjects experienced similar relative hypertrophic responses and both significantly increased strength along the same timeline. Of note is that the vast majority of these studies used untrained subjects. Moreover, the Position Paper called for studies that more closely resembled the training of female athletes. Some studies in this area trained only one joint or used only a handful of exercises. Most extensively incorporated resistance training machines, rather than free weights. The general theme with this type of research was simply laying the groundwork, and exploring specific responses of female muscle to resistance training protocols. Once it was established that women responded in a similar fashion to men, investigators began to focus on training methodology. The next broad area of inquiry dealt with varying the methods of training to explore how that might affect outcomes, like hypertrophy, strength, or power.

**Strength Training Methodology in Women**

Regarding training methodology, the research in this area covers a variety of topics ranging from the effects of varying sets and repetitions, comparing machines and free weights, total body versus split programs, comparing eccentric and concentric protocols, and evaluating different periodization schemes. Coming out of the “machine age” of the 1980s (Nelson, 1982) in which resistance training machines like Nautilus became prominently associated with strength training, Brian Boyer (1990) compared the effects of machine and free weight training programs on female subjects. Published the year after the Position Paper, the study recruited untrained women to train 3 times per week for 12 weeks using either free weights, Nautilus machines, or a Soloflex machine. Loads ranged from 6 to 10RM, with all three groups demonstrating increased strength. In a nod to the principle of specificity, Boyer noted that participants showed better strength improvements on the equipment on which they trained.

Stone and Coulter (1994) examined the effects of different repetition ranges on muscular strength and endurance using untrained female subjects. For 9 weeks different groups of subjects performed either 3 sets of their 6-8RM, 2 with their 15-20RM, or 1 set with their 30-40RM. Using a combination of free weights and machines for the training program, the authors found that the heavier training was more effective at inducing strength gains, while the lighter and higher-repetition training
produced better results for muscular endurance. Using a similar exercise protocol to Staron’s previous work, Rana and colleagues (Rana et al., 2008) compared the effects of low-velocity training, in which each repetition lasted from 14-20 seconds to a traditional program using 6-10RM and an endurance protocol that used a 20-30RM in 3 lower-body exercises. While the low-velocity program improved both strength and endurance in the untrained women, it was not as effective as either the strength- or endurance-specific program at enhancing those variables. Ben-Sira and colleagues (Ben-Sira, Avalon, & Tavi, 1995) used untrained female subjects to compare supra-maximal eccentric training on knee extension exercises to concentric, eccentric, or conventional training over the course of 8 weeks. There were no differences in strength measures between the four protocols at the end of the program, nor was there a significant increase in thigh girth. The lack of measurable hypertrophy was likely due to the short duration of the program. Nonetheless, the authors pointed to the lack of muscular growth in their conclusion, assuring women that they need not worry about developing excessive muscularity, at least in the short term.

Calder and colleagues (Calder, Chilibeck, Weber, & Sale, 1994) used resistance training machines to compare a total body to a “split” strength program. Untrained women exercised on the weight stack machines twice weekly for the total body group or 4 times weekly for the split group, for a total of 20 weeks. Loads ranged from 6-10RM with both groups demonstrating similar increases hypertrophy, strength, and decreased skinfold thickness at the end of the program. While this study held volume constant, several studies in the early 2000s sought to examine the effect of training volume on muscular size, strength and power. This line of inquiry was also tied to the “machine age” insofar as the creator of Nautilus machines, Arthur Jones, was a fervent advocate of a training system that called for one set to failure.

Sanborn and her colleagues (Sanborn et al., 2000) from Appalachian State, including Meg Stone, a member of the Women’s Committee, cited Jones’ theory in the introduction of their work comparing the effects of single versus multiple sets. Unlike previous investigations, this study exclusively utilized free weights, including variants on the Olympic lifts, and was primarily interested in performance outcomes, rather than hypertrophy. Using squats, straight-legged deadlifts, bench press, mid-thigh pulls and more for 8 weeks, researchers compared the effect of one set to failure between 8-12 repetitions to three sets ranging from 2-10 repetitions. The previously untrained subjects in the multi-set group increased both their 1-RM strength in the squat and counter-movement jump height more than those who performed only one set. The following year, Schlumberger and colleagues (Schlumberger, Stec, & Schmidtbleidher, 2001) compared the effects of one set of 6-9 repetitions to 3 sets of 6-9 repetitions in
women with at least six months of training experience. Using exclusively weight machines, the women trained twice weekly for 6 weeks. Women in the 3 set protocol demonstrated superior strength increases to those in the 1 set group, demonstrating that maximizing strength requires higher volume.

The mid-2000s saw a proliferation of the effects of periodized strength training on female subjects. Prior studies generally used “Progressive Resistance Exercise (PRE)” protocols in which weight was increased once participants could easily perform the prescribed number of repetitions. Periodized protocols varied the intensity of the resistance along with the volume in a variety of ways and sought to compare which was optimal for inducing hypertrophy or various parameters of performance. Kraemer and colleagues (Kraemer et al., 2004) applied a periodized program to four groups of untrained women over the course of 6 months. The women were in either total-body or upper-body groups that focused on either hypertrophy or power. Repetitions ranged from 8-12 in the hypertrophy group and 3-8 in the power group with subjects increasing the amount of weight lifted over the course of 3 months as the repetitions dropped. Both groups started over at the higher repetition range at the end of three months and again slowly increased weight, thus performing two mesocycles. The women performed a combination of free weight exercises like squats and bench press as well as pulley exercises like lat pulldowns. All groups increased the size, strength and power of trained musculature with the heavier protocol showing more consistent increases in muscle size. In addition to the efficacy of heavy training, this article also demonstrated continued improvement in those variables at the conclusion of this study, whereas studies using PRE protocols typically observed a plateau in those same measures after 3 to 4 months of continuous training (Herrick & Stone, 1996).

Kok and colleagues (Kok, Hamer, & Bishop, 2009) compared different periodization models, specifically linear versus undulating in recreationally active women. Whereas linear periodization increases intensity and decreases volume on a weekly basis, daily undulating periodization changes both between training sessions. Using recreationally active women who were inexperienced with strength training, the authors found that 9 weeks of free weight and machine training produced similar results with both periodization models. Both training programs increased muscular strength, size and power. Using sedentary women and high repetition ranges, between 15-30, de Lima and colleagues (de Lima et al., 2012) found linear periodization to be superior at improving muscular strength and body composition, while daily undulating was more effective at increasing muscular endurance. Prestes and colleagues (Prestes, de Lima, Frollini, Donatto, & Conte, 2009) compared linear to reverse linear periodization for their effectiveness at altering strength and body composition. Reverse linear periodization entails starting at with heavy loads and decreasing them while increasing volume as the
training progresses. Women with a minimum of 6 months of strength training experience used a combination of free weights and machines to lift loads ranging from 4-14 repetitions over the course of 12 weeks. While both groups increased strength, linear periodization appeared to be superior at both increasing muscular strength and size.

Bartolomei and colleagues (Bartolomei, Stout, Fukuda, Hoffman, & Merni, 2015) compared block periodization, in which a “block” of training is focused on developing a specific parameter like hypertrophy or muscular power, to a weekly undulating periodization program. A group of women with at least two years of recreational strength training experience performed a variety of power lifts and machine exercises with repetitions ranging from 3-10. At the end of the 10 week intervention, weekly undulating periodization was more effective at increasing muscular size and strength than the block model. Robert Kell (2011) compared the effects of linear periodization on trained male and female recreational athletes. Though experienced with strength training, none of the athletes had used periodized programs. At the end of 12 weeks of training that included machines and free weights, both male and female subjects significantly increased their strength with women showing a greater relative increase than their male counterparts.

As researchers in the 1990s and 2000s began to tinker with program design and examine the effects on precursors to athletic performance, like muscular strength, size, or power, other work investigated the effects of strength training on variables more closely related to athletic performance. A concurrent line of inquiry to studies on training methodology was that of applied programs for athletes. In that work, researchers implemented training programs for competitive athletes to see not only how they might affect muscular strength and power, but also how they might affect time trial performance in runners, serve or throwing velocity, running economy and more.

**Applied Programs for Athletes**

Shortly after publication of the Position Paper, Fry and colleagues (Fry et al., 1991) studied the effects of an off-season strength program on NCAA Division I volleyball players. The 12-week program included free weights, variants on the Olympic lifts, machine exercises and a plyometric component. As a result of the training, athletes significantly increased fat-free mass, increased 1RM strength, increased shoulder flexibility, improved their vertical jump, and decreased their 2-mile run time. Somewhat surprisingly, sprint performance did not improve, and vertical jump endurance declined. After noting in their discussion that the athletes had increased muscle mass, the authors went on to point out that “for players concerned about muscle hypertrophy” strength improvements could largely be accounted for by factors other than muscle size (Fry et al., 1991, p. 181).
In a less comprehensive intervention, Hoff and Almasbakk (1995) had competitive female handball players perform throwing practice combined with the bench press exercise for 9 weeks using 3 sets of 5-6 repetitions. The weight trained group improved their throwing velocity more than a throwing-only control, with a significant improvement in velocity with a 3 step run-in. In another one-exercise intervention Hoff and colleagues (Hoff, Helgerud, & Wisloff, 1999) had competitive female cross country skiers perform 3 sets of 6RM with a pulley exercise designed to simulate the double-poling movement. After 9 weeks of training, the experimental group improved its time to exhaustion nearly 80% more than a control that performed a non-specific low-intensity strength program. Heavy pulley training decreased time to peak force production and improved skiing economy. Bishop and colleagues (Bishop, Jenkins, Mackinnon, McEniery, & Carey, 1999) added five sets of squats to failure on a “Plyopower” machine to the regimen of experienced female cyclists. The study incorporated heavy loads, with failure reached between 2-8 repetitions. While the 12-week intervention increased 1RM strength, it did not produce improvement in a one-hour time trial.

Petko and Hunter (1997) tracked the effects of basketball workouts and a sport-specific conditioning program across the playing careers of NCAA Division I women’s players. The athletes significantly improved their strength and vertical jump throughout their time in the program though, as the authors observed, increases in strength were not matched by increased fat-free mass. In keeping with other work, the female athletes showed a greater relative increase in strength than male counterparts who participated in a comparable program. Consistent with the interest in the variable of training volume in the early 2000s, Kraemer and colleagues (Kraemer et al., 2000) studied the impact of single versus multiple sets on strength, power, body composition and serve velocity in female collegiate tennis players. Both groups performed a variety of barbell, dumbbell, and machine exercises, with the single-set group performing their 8-10RM on each. The multi-set group performed repetitions ranging from 4-15 for 2-4 sets in a linearly periodized program. At the end of the 9 month intervention, the higher volume program more effectively enhanced muscle size, strength, serve velocity, and lean mass relative to the single-set program. From a performance standpoint it is noteworthy that the single-set group did not show a significant improvement in serve velocity. In a similar study published 3 years later, Kraemer and colleagues (Kraemer et al., 2003) showed that a 9-month program of periodized strength training was more effective at increasing vertical jump height, serve, forehand, and backhand ball velocities than a non-periodized program in collegiate female tennis players.

Kelly and colleagues (Kelly, Burnett, & Newton, 2008) added heavy strength training to the workouts of recreational endurance runners. The women performed a total-body strength program
using both machines and free weights for 3 sets of 5 repetitions. When evaluated against a running-only control after 10 weeks, the strength trained group showed a “statistically non-significant, but practically important” improvement in their 3 kilometer running performance (Kelly, Burnett, & Newton, 2008, p. 402). In a similar study, Johnston and colleagues (Johnson, Quinn, Kertzer, & Vroman, 1997) investigated the impact of a total-body strength program in experienced female endurance runners. Relative to a running-only control, the strength trained group significantly improved their running economy. Agustsson and colleagues (Augustsson, et al., 2011) compared the strength and vertical jump performance of high school-aged female volleyball players who were given either a general and unsupervised strength program or provided with supervision and an individualized program. Athletes who were supervised and had their programs tailored to their needs improved all strength measures, but not vertical jump, more than unsupervised counterparts, highlighting the potential importance of strength coaches.

Citing a lack of research on female athletes generally and female soccer players in particular by 2012, Grieco and colleagues (Grieco, Cortes, Greska, Lucci, & Onate, 2012) examined the effects of 10-week strength program on running economy and aerobic capacity in Division I soccer players. Using an advanced and sport-specific program the intervention included both barbell and dumbbell variants on Olympic lifts, plyometrics, medicine ball and bodyweight exercises performed for 3 sets of 4-12 repetitions. Peak oxygen consumption significantly increased following the intervention with equivocal impact on running economy. Noting a lack of research on the efficacy of Olympic variants on performance variables in female athletes, Ayers and colleagues (Ayers, DeBeliso, Sevne, & Adams, 2016) compared the effectiveness of hang cleans and hang snatches at improving vertical jump height, sprint speed, and squat strength. The NCAA Division I volleyball and softball players involved in this study performed the exercises for 5 sets of 3 repetitions as part of a more comprehensive program. Following the 6-week intervention both groups similarly and significantly improved their vertical jump, 40 yard sprint time, and 1RM squat. Given that the vast majority of each group’s program was the same, with the exception of one exercise, the more important takeaway is that the comprehensive program was effective at enhancing several important performance parameters.

In addition to strength programs for team sport athletes, several recent studies have looked at the impact of strength training on endurance performance. Skattebo and colleagues (Skattebo, Hallen, Ronnestad, & Losnegard, 2016) added 3 sets of 4-10 repetitions of lat pulldowns, a pulley exercise intended to mimic double-poling, and tricep pressdowns to the program of female junior cross-country skiers. The intervention group, along with the control, also performed a variety of core and bodyweight
exercises. Following a 10 week intervention, the group with additional strength training had increased arm size and muscular strength but did not differ significantly from the control group in peak aerobic capacity, power output on an ergometer, or submaximal oxygen cost. Vikmoen and colleagues (Vikmoen et al., 2016) added lower body strength training to the workouts of women experienced in both running and cycling. Following an 11 week intervention that included 3 sets of 4-10RM on smith machine half squats, leg press, resisted hip flexion and calf raises, the training group improved 1RM strength, jump height, and increased the size of the trained musculature. The training group did not, however, differ after the intervention from a group that performed only endurance training in measures of aerobic capacity, running economy, or a 40-minute all-out run. The following year, Vikmoen and colleagues (Vikmoen, Ronnestad, Ellefsen, & Taastad, 2017) published a similar study, this time testing the effects of strength training on all-out performance following prolonged submaximal cycling and running. The group that performed both strength and endurance training improved their all-out performance while the endurance-only group did not. Additionally, the strength-trained subjects demonstrated reduced oxygen consumption and heart rates during the prolonged exercise, whereas the endurance group did not.

**Hormonal Effects of Strength Training in Women**

A handful of studies investigated endocrine, rather than musculoskeletal effects of strength training in women. Kraemer and colleagues (Kraemer et al., 1991) used two different strength training protocols to examine hormone responses in recreationally trained men and women. Subjects performed 5 sets of their 5RM on 8 different barbell, pulley, and machine exercises with long rest intervals on one day, and 3 sets of their 10RM with shorter rest intervals on another in random order. Pre- and post-exercise plasma hormone levels were measured for both protocols. Female subjects had higher initial growth hormone (hGH) levels than males and saw a marked increase in hGH levels following the higher repetition protocol. No effect was observed for the heavy protocol on hGH levels in women, nor was testosterone affected by either protocol. In a similar study published two years later, Kraemer and colleagues (Kraemer et al., 1993) mixed the loads and rest periods into six different combinations of 5RM and 1-3 minutes of rest or 10RM with 1-3 minutes of rest. All of the recreationally trained women performed each workout over the course of the study. As with the prior study, testosterone was unaffected by program alterations and hGH increased after the higher repetition, shorter rest combination. Cortisol showed a significant increase with the same protocol.

Hakkinen and colleagues (Hakkinen, Pakarinen, & Kallinin, 1992) trained women for 3 weeks using heavy loads ranging from 1-3RM and 5-10RM on squats, leg press, and leg extensions. The
previously untrained subjects increased their muscular size and strength, however, there were no significant changes in testosterone, cortisol, or sex hormone binding globulin following the intervention. Using the same protocol that Staron used to investigate fiber type changes and hypertrophic adaptations to strength training, Kraemer and colleagues (Kraemer et al., 1998) studied hormonal responses in men and women. Subjects trained for 8 weeks between 6-12RM on squats, leg press, and leg extensions. Unlike previous work, female subjects demonstrated increased pre-exercise testosterone after 6 and 8 weeks. Sex hormone binding globulin, which affects testosterone use, was also increased in women at the study’s conclusion.

**Bone Density and Injury Prevention**

Following the announcement of the Women’s Health Initiative and a directive to include female subjects in research, the 1990s saw a raft of studies examining the effects of strength training on bone density and osteoporosis in women. By the end of the decade dozens of articles demonstrated a positive relationship between strength training and bone density (Layne & Nelson, 1999). Randomized clinical trials indicated that both high intensity endurance and strength training were capable of mitigating and possibly reversing bone loss in pre- and post-menopausal women (Wolff, van Croonenborg, Kemper, Kostense, & Twisk, 1999). The data were sufficiently compelling for the International Osteoporosis Foundation to include resistance exercise in its exercise recommendations for adults and the elderly (“Exercise,” n.d.). The high volume of research in this area makes it important to mention, though a detailed discussion is beyond the scope of this work, as it is not closely related to female athletes.

Strength training has been used to prevent injuries in female athletes, however, and this has been another area of extensive research. Physicians began commenting on the rate of knee injuries in female athletes in the late 1970s; it was not until the mid-1990s, however, that researchers began to focus on the disparity in anterior cruciate ligament (ACL) injuries in female relative to male athletes (Eisenberg & Allen, 1978). Female athletes, the research showed, were 2-3 times more likely to rupture their ACLs than male athletes playing the same sport (Arendt & Dick, 1995; Agel, Arendt, & Bershadsky, 2005; Prodromos, Han, Rogowski, Joyce, & Shi, 2007). In response, researchers began to investigate the disparity, citing possible causative factors ranging from femoral intercondylar notch size, to hormones, poor footwear, muscular imbalances or weakness, and innate limb alignment (Souryal & Freeman, 1993; Ireland, Guadette, & Crook 1997; Wojtys, Huston, Lindenfeld, Hewett, & Greenfield, 1998; Lephart, Ferris, Riemann, Myers, & Fu, 2002). Athletic trainers, physical therapists, and physicians began to devise exercise intervention programs to remediate some of the correctable factors, like muscle
strength or landing technique (Hewett, Lindenfeld, Riccobene, & Noyes, 1999; Myer, Ford, & Hewett, 2004). While these programs incorporated strength training, they typically incorporated relatively light resistance like exercise bands or static exercises like abdominal bridges (Herman et al., 2008). Other common components include a focus on plyometrics and landing from single and double-leg jumps while minimizing medial knee movement. Two recent reviews have found favorable results, with prevention programs reducing the risk of ACL ruptures by as much as 52% in female athletes (Sadoghi, von Keudell, & Vavken, 2012; Taylor, Waxman, Richter, & Shultz, 2015). Other research has shown mixed results. With only 2 of 10 studies analyzed demonstrating a significant decrease in ACL injuries, Stevenson and colleagues (Stevenson, Beattie, Schwartz, & Busconi, 2014) declined to recommend general implementation of ACL prevention programs. As they acknowledged, however, the cumulative results point toward a trend in injury reduction. Key factors appear to be program design, with programs that include multiple training modalities being more effective, and timing of the intervention, with those that include both pre-season and in-season components showing better results.

As more and more researchers implemented ACL injury prevention programs, some began to investigate whether they might actually enhance performance. Citing poor compliance with the injury prevention programs, researchers sought to test whether modified versions of the programs could increase vertical jump height, sprint speed, or aerobic power output (Noyes, Barber-Westin, Smith, & Campbell, 2013). If so, they reasoned, it would provide an additional selling point for the programs and help improve buy-in from athletes, coaches, and parents. Hewett and colleagues devised the “Sportsmetrics” ACL injury prevention program in 1996, with the earliest iteration of the program including barbell and machine training in addition to plyometrics and jumping exercises (Hewett, Stroupe, Nance, & Noyes, 1996). A revised version of the program that included resistance band and bodyweight exercises as the strength training component was employed on adolescent female athletes in a variety of sports (Barber-Westin & Noyes, 2017). Following 6 weeks of training, high school volleyball players improved knee alignment when landing from a jump, increased their vertical jump height, had stronger abdominal musculature, and improved their estimated aerobic capacity (Noyes, Barber-Westin, Smith, & Campbell, 2011). Female high school basketball players significantly increased their estimated aerobic capacity and vertical jump, with no effect on sprint performance (Noyes, Barber-Westin, Smith, Campbell, & Garrison, 2012). Likewise, high school soccer players significantly increased their estimated aerobic capacity and vertical jump, as well as their agility and sprint speed (Noyes, Barber-Westin, Smith, & Campbell, 2013).
Beyond knee injuries, one study attempted to quantify the differences in injuries related to weightlifting between men and women. Using hospital data from 2002-2005, Quatman and colleagues (Quatman, Myer, Khoury, Wall, & Hewett, 2009) found that women were more likely to suffer accidental injuries in the weight room. These might include tripping over weights or machines, or having weights fall on them. Men, on the other hand, were more likely to report injuries related to over-exertion or poor form like strains or sprains. The authors speculated that the higher rate of accidental injuries in women may be due to a combination of little or untrained supervision, or lack of education on training techniques.

CONCLUSION

As we approach the 30th anniversary of the publication of the NSCA’s Position Paper on strength training for female athletes, what do we know? Some of the things that were known at the time of the Paper’s publication have been validated by additional work. As an example, the Women’s Committee noted, and further research confirmed, that men and women showed comparable hypertrophy in response to strength training, relative to initial muscle mass (Staron et al., 1989; Staron et al., 1994). Several authors were eager to point out, however, that significant strength gains can be made by female athletes without substantial hypertrophy (Ben-Sira et al., 1995). Work in the years immediately after the Paper demonstrated that, not only do men and women have a similar muscle fiber type distribution, but those fiber types shift toward type IIa in both sexes following resistance training (Miller et al., 1993; Staron et al., 1994). Compared to men, women are likely to see a larger relative increase in muscle strength (O’Hagan et al., 1995; Hubal et al., 2005). Multiple sets and periodized programs are more effective than single sets or non-periodized programs at increasing muscle strength, size, and power (Sanborn et al., 2000; Schlumberger et al., 2001; Kraemer et al., 2004).

Results are mixed regarding combined strength and endurance training. The few studies conducted examining the effects of the combination on muscle strength have shown no deleterious effects of endurance exercise on strength performance (Volpe et al., 1993; Gravelle & Blessing, 2000). While there is some evidence that strength training can improve the economy of movement, the verdict is far from unanimous (Johnston et al., 1997; Hoff et al., 1999; Skattebo 2016; Vikmoen, 2016). It should be noted that the studies that examined these parameters typically included only a few exercises either for the upper or lower body exclusively. At the very least, while strength training may not improve endurance performance, it does not appear to hamper it either.

Authors of the Position Paper called for programs that mimicked the actual modes of training used by female athletes. Few studies have heeded this call. Most studies used untrained subjects, and
many used protocols in which exercise machines accounted for half or more of the exercises. Recent studies have moved toward programs that are likely to be more similar to what collegiate and elite athletes actually perform, including variants on the Olympic lifts, free weights, training of the core musculature, and training the whole body (Grieco et al., 2012; Ayers et al., 2016). That said, one of the most glaring gaps in knowledge is how many female athletes actually strength train and what exercises they are likely to perform. As noted by Quatman and colleagues (Quatman et al., 2009) there are no current epidemiology studies or databases that describe female and male participation in strength training as part of their sport preparation. That work included statistics on male and female participants in the competitive sport of Olympic weightlifting as a proxy of participants in strength training as a whole, for lack of better data. Research surveying the number of male and female athletes who strength train for sports and the methods employed would be useful.

Another potential direction for future research is the examination of the influence of a strength program on sport-specific or in-game performance measures. Many studies have measured components of athletic performance, like muscle strength, vertical jump height, and aerobic capacity. A few studies have measured closer proxies of sport performance like throwing or serving velocity. Research in the coming years could take advantage of GPS and accelerometer technology to examine the relationship between muscular strength and endurance and the maintenance of running velocity during a women’s soccer match or maintenance of vertical jump height throughout a volleyball match. Using game statistics, researchers could analyze correlation between strength and power of the core musculature and slugging percentage in softball. Just as Noyes and colleagues (Noyes et al., 2012; Noyes et al., 2013) sought to demonstrate performance enhancement from injury-prevention programs, research directly connecting strength training and sport performance may prove helpful in encouraging female athletes to lift weights.

Beyond physiology, the Women’s Committee cited the importance of female role models in encouraging female athletes to initiate and continue strength training (Holloway et al., 1989b). According to Laskowski and Ebben (2016), women hold approximately 32% of strength coaching positions at NCAA Division I institutions. While this is a far cry from the 3 who held such positions in 1991, in light of the fact that females comprise nearly 45% of NCAA athletes, it appears that women are still underrepresented as strength coaches (National Collegiate Athletic Association, 2016). Of those who are strength coaches, surveys have shown that many believe that their careers are hampered by a “glass ceiling” and that their opportunities to move beyond the level of assistant coach or into administration are limited (Massey & Vincent, 2013; Laskowski & Ebben, 2016). Other commonly-cited
drawbacks include a lack of respect from players, coaches, and athletic administrators, and not being able to work with male teams. Indeed most female strength coaches do work with female athletes, and there is some evidence that male athletes are resistant to working with female strength coaches (Magnusen & Rhea, 2009). As Magnusen and Rhea (2009) point out, this resistance on the part of male athletes toward female strength coaches makes it all the more important that both male and female athletes be exposed to female strength role models early in their athletic careers. In spite of these obstacles nearly all female strength coaches report high job satisfaction (Massey & Vincent, 2013; Laskowski & Ebben, 2016).

While some female strength coaches believe that their career trajectory is limited by their gender, sociologist Shari Dworkin (2001) has discussed another gendered limitation in the weight room. Dworkin described a “glass ceiling” on female muscularity in which women consciously work to avoid becoming what they perceive to be ‘too big’ as a result of their training. In Western culture, one of the key markers of masculinity is muscularity. The ideal masculine body is one with large and well-defined muscles, implying aggression and the ability to physically dominate. Conversely, the feminine ideal has been described as the opposite; a body that is beautiful, small, lean, and weak (Roth & Basow, 2004). When women express concerns about getting ‘too big,’ then, they are often implicitly expressing a concern about transgressing gender norms. Through interviews at fitness centers Dworkin described the notion, held by almost three-quarters of women who lifted weights, of a desirable upper limit on both their own strength and muscularity. The women often pointed to the physiques of female bodybuilders and expressed “explicit fear and repulsion” of developing that aesthetic (Dworkin, 2001, p. 337). Many of the subjects avoided lifting altogether in favor of exclusively cardiovascular exercise while others described lifting light weights in order to avoid “bulking up.” Of those who did lift regularly, reports of not increasing the weight, decreasing weight, or simply abandoning lifting entirely for a time if they felt they had gotten too big, emerged.

Though Dworkin’s work was with recreational strength trainers, there is evidence that collegiate athletes harbor many of the same beliefs. Prior to the dramatic increase in the number of female athletes in the 1970s, femininity and athleticism were generally held to be in conflict (Heywood & Dworkin, 2003). Sports were thought to develop traditionally masculine qualities like aggression, competitiveness, and physical power. The female athlete, then, was considered to be at risk of developing masculine qualities (Cahn, 2003). While associations between athleticism and masculinity have weakened, the association between muscularity and masculinity largely have not. As such, female athletes have been said to have to navigate multiple contexts, including the athletic and the social, and
may develop multiple understandings of the muscularity developed through sports and weight training (Cox and Thompson, 2000; George, 2005). As an example, a survey of 14 female NCAA Division I athletes from 7 different sports found that the athletes entertained conflicting perceptions about their own muscularity (Roth & Knapp, 2017). While they viewed their muscles with pride as a symbol of their dedication and strong work ethic, all identified a point at which that muscularity would be excessive. To avoid becoming too muscular and transgressing gender and preferred aesthetic norms, the athletes described “holding back” their effort and the weight lifted during their workouts. Others did extra cardiovascular work, unbeknownst to their strength coaches, to try to minimize hypertrophy induced by the strength program. This behavior also was noted in other work on female NCAA Division I soccer players, with some even describing wearing clothes to cover their muscles when outside of the athletic environment (George, 2005). Competitive female adolescent track athletes also expressed reservations about becoming overly “bulky” as well as concerns about perceptions of their muscularity outside of the competitive environment (Mosewich, Vangool, Kowalski, & McHugh, 2009). Despite fears of excessive muscularity, however, many female athletes have embraced the weight room and said they both enjoyed lifting and believed that it would help their performance (Roth & Knapp, 2017).

While more is known about the effects of strength training on female athletes than at the time of the Position Paper, there is still more to be learned. It is worth pointing out that many contemporary strength training studies do include both male and female subjects. One could argue that this in itself is progress because both sexes are being treated equally by enduring the same rigorous training protocols. There is much truth to this. Given that differences in female anatomy and physiology have been described as potential factors in injuries ranging from ACL tears, to stress fractures, and concussions, however, it seems prudent to advocate for continued female-specific research or reporting of results (Wojtyś, et al., 1998; Kelsey et al., 2007; Wunderlie, Hoeger, Wasserman, & Bazarian, 2014). Nonetheless, given what is known about adaptations to strength training programs by female athletes, the basic conclusion of the Position Paper still holds true nearly 30 years later, “it appears that males and females should train for strength in the same basic way, employing similar methodologies, programs, and types of exercises” (Holloway et al., 1989).

REFERENCES


Kraemer, W.J., Nindl, B.C., Ratamess, N.A., Gotschalk, L.A., Volek, J.S., Fleck,


doi: 10.1136/bjsports-2013-092358

Timeline: America’s war on drugs. (2007). National Public Radio, NPR.org,


doi: 10.1371/journal.pone.0150799


doi: 10.1177/03635465980260050301

doi: 10.1007/s001980050

doi: 10.1097/HTR.00000000