

Original Article

Does a Validated Circum-menarcheal Arm Bone-Loading Index Predict Post-menarcheal Muscle Function in Adolescent Females?

Brian Josephson¹, Timothy Chan¹, Carol A. Sames², Tamara A. Scerpella^{1,3}, Jodi N. Dowthwaite^{1,3}

¹Department of Orthopedic Surgery, SUNY Upstate Medical University, Syracuse, NY, USA;

²College of Health Professions, SUNY Upstate Medical University, Syracuse, NY, USA;

³Department of Orthopedics & Rehabilitation, University of Wisconsin-Madison, Madison, WI, USA

Abstract

Objective: In this exploratory analysis, we evaluated our validated 3-year circum-menarcheal arm bone loading index (ARMBLI) as a predictor of upper extremity muscle function indices in post-menarcheal females. **Methods:** Forty-eight participants met inclusion criteria for non-dominant arm analyses, 1 to 2.5 years post-menarche: 1) menarche date (GYNAGE: yrs, menarche-centered age); 2) ARMBLI (3-yrs prior loading exposure); 2) grip strength (kg; primary outcome), 3) Biodex peak torques (Nm: shoulder, elbow; flexion, extension; secondary outcomes), 4) DXA arm non-bone fat-free mass (ARMNBFFM, kg), 5) anthropometry [ulna length (mm), height (cm), weight (kg)]. Multiple linear regression evaluated muscle function in the total sample (n=48), gymnasts (n=24), and non-gymnasts (n=24), accounting for key factors: 1) whole-body model (GYNAGE, height, weight, ARMBLI); 2) hypertrophy model (GYNAGE, ulna length, ARMNBFFM, ARMBLI). **Results:** After accounting for key factors, ARMBLI was an independent negative predictor of grip strength in the total sample [hypertrophy model B= -0.202, p=0.029] and non-gymnast subsample [hypertrophy model B= -0.935, p=0.030; whole-body model B= -1.127, p=0.005]. Conversely, ARMBLI was an independent positive predictor of shoulder flexion peak torque in the total sample [whole-body model B= +0.270, p=0.036]. **Conclusions:** Despite prior work showing ARMBLI predicts ARMNBFFM, ARMBLI showed inconsistent predictive value for muscle function indices.

Keywords: Adolescent, Exercise, Female, Grip Strength, Peak Torque

Introduction

From birth to death, skeletal muscle mass, coordination, and power are required for health and function. Most muscle mass is acquired during childhood and adolescence, peaking age 20 to 30 years due to further hypertrophy¹. Muscle hypertrophy plays a direct role in muscle function, assessed via gains in non-bone appendicular lean mass or muscle cross-sectional area (CSA), but neurological

mechanisms such as enhanced motor neuron recruitment can boost muscle functional capacity without significant hypertrophy². Accordingly, local muscle CSA does not fully explain functional variance, but related research underrepresents females and fails to evaluate the influence of physical maturity^{3,4}. Additional functional variance is explained by youth neuromuscular properties that may set potential capacity for muscle gains across the lifespan⁴⁻⁶, by improving necessary coordination between neurons and muscles before early adult hypertrophy⁵. As baseline adult neuromuscular function, muscle mass, and strength is established during childhood and adolescence, the importance of boosting youth physical activity has been highlighted across platforms⁷⁻⁹. Healthy youths are advised to sample varied activities, but adolescent data are lacking on how exposures to contrasting loading types influence muscle function around menarche – the critical period of rapid lean mass hypertrophy in females.

The authors have no conflict of interest.

Corresponding author: Jodi N. Dowthwaite, PhD, Department of Orthopedic Surgery, 750 E. Adams Street, Syracuse, NY 13210
E-mail: dowthwaj@upstate.edu

Edited by: G. Lyritis

Accepted 5 February 2026



Beyond empowering movement during youth, adolescent muscle development sets the foundation for adult fitness and well-being. Better adolescent muscle function is linked to lower adiposity and metabolic risk factors, with wide-ranging youth health benefits¹⁰, but also persistently lower adult BMI, adiposity, triglyceride levels, and insulin resistance¹¹. Grip strength is an important functional metric across the lifespan, as it predicts overall health, disease status, fracture risk, and mortality in adulthood¹². Alarming, in a longitudinal cohort of 623 males and females, compared to stronger youths, those in the lowest tertile for adolescent grip strength and shoulder extension/flexion performance were ≥ 4 times as likely to perform in the lowest tertile in adulthood¹³. Associated risks are modifiable, as vigorous adolescent physical activity and resistance training may yield dose-dependent benefits in muscle function, with low risk of injury when properly supervised^{14,15}. Because adolescent exposures may influence long-term health, we must learn how youth activity affects later muscle function, particularly grip strength in post-menarcheal females – an understudied, vulnerable group.

Many studies have monitored peak torques and grip strength in youths affected by traumatic injuries and chronic conditions^{16,17}, but grip strength is rarely evaluated in healthy, active girls. In U.S. 12-15 year olds, including girls, reported resistance training (≥ 2 sessions per week) was linked to benefits in grip strength “muscle quality” (grip strength per kg non-bone lean mass), but physical maturity and doses of other activities were not directly evaluated¹⁸. Most muscle function studies in healthy youths assess short-term interventions or performance in specialized athletes, often including elites mixed with adults, rarely evaluating sport-specific training dose^{3,19-28}. Most assessed males alone or compared males to small female samples, without evaluating sex-specific maturity variation^{3,23,26,27,29}. Most focused on lower extremity training gains in lean mass or muscle CSA rather than functional outcomes. New Portuguese studies have evaluated physical maturity as a factor in upper extremity muscle function, including females^{28,30}, but neither assessed rapid hypertrophy around menarche or evaluated the role of chronic activity dose. In a third Portuguese cohort analysis of 5 to 11 year olds (unspecified maturity), annual hip accelerometry (1-week per year) did not predict grip strength development over 3 years³¹. Short-term objective assessments of lower body movement in youths present limitations for predicting upper extremity functional development: 1) activity varies over extended time courses; 2) contrasting loading modalities are site- and sport-specific (e.g. swimming, cycling, long-distance running, soccer, gymnastics); 3) hormonal milieu and body size change across time. In adolescent girls, the role of maturity- and site-specific loading doses on upper extremity muscle function remains unclear.

To alleviate high risk of sarcopenia and fragility fractures in women, we must understand how long-term youth activity patterns shape female musculoskeletal outcomes in late adolescence – the foundation for adult function

and risk. Researchers evaluate pediatric physical activity and/or sedentary behavior as developmental factors for the functional muscle-bone unit; they correlate loading patterns with youth bone properties, accounting for non-bone lean mass, muscle CSA, or muscle function indices, including females of normal, overweight, and obese weight status^{32,33,34}. Positive relationships have been reported between muscle function indices and bone properties^{6,17}. Yet, most only report bone outcomes adjusted for muscle covariates, missing the opportunity to evaluate magnitude, direction, and significance of correlations between activity doses and muscle function metrics^{35,36}. A key knowledge gap remains on whether diverse loading exposures shape development of female muscle function. It is particularly important to evaluate the circum-menarcheal period when bone accrual is highest and most rapid³⁷.

In prior research, our team laid groundwork to close this knowledge gap, showing that osteogenic exercise around menarche predicts bone and muscle indices, evaluating 3 years that include peak velocity for bone accrual³⁸. That work validated our circum-menarcheal, non-dominant arm bone-loading index (ARMBLI) in 2 cohorts from our longitudinal study of sport-specific musculoskeletal development³⁸. ARMBLI evaluated training doses of >50 diverse physical activities, incorporating sport-specific magnitude, frequency, and velocity (impact) of loading, summing total hours for all physical activities reported over 3-years circum-menarche³⁸. These cohorts included girls who participated in artistic gymnastics – a highly osteogenic sport – with varying exposures. ARMBLI clearly quantified maturity-specific osteogenic loading, as it was a significant, positive predictor of post-menarcheal radius bone mass and density, accounting for gynecological age at measurement (GYNAGE: menarche-centered age)³⁸. ARMBLI was also a significant, positive predictor of local muscle mass, explaining roughly 20% of independent variance in non-dominant arm non-bone lean mass (ARMBLI_{FFM})³⁸. Thus, high site-specific bone-loading activity circum-menarche (ARMBLI) predicted high arm muscle mass post-menarche.

In the current exploratory analysis, we set out to address the focal knowledge gap, asking: Does circum-menarcheal arm loading predict local muscle function in post-menarcheal girls? We hypothesized that ARMBLI would predict higher grip strength and peak torques, for both shoulder and elbow, in girls exposed to heterogeneous osteogenic loading via diverse physical activity types around menarche. We aimed to: 1) account for physical maturity and body size; 2) account for maturity, local lever length, and muscle hypertrophy; 3) evaluate whether observed patterns held across the total cohort, “gymnast” subsample, and “non-gymnast” subsample. Grip strength is our primary outcome, as it is highly valued as an indicator of lifespan muscle function and health. We provide Biodex peak torque results as secondary outcomes, emphasizing effect sizes to inform future research.

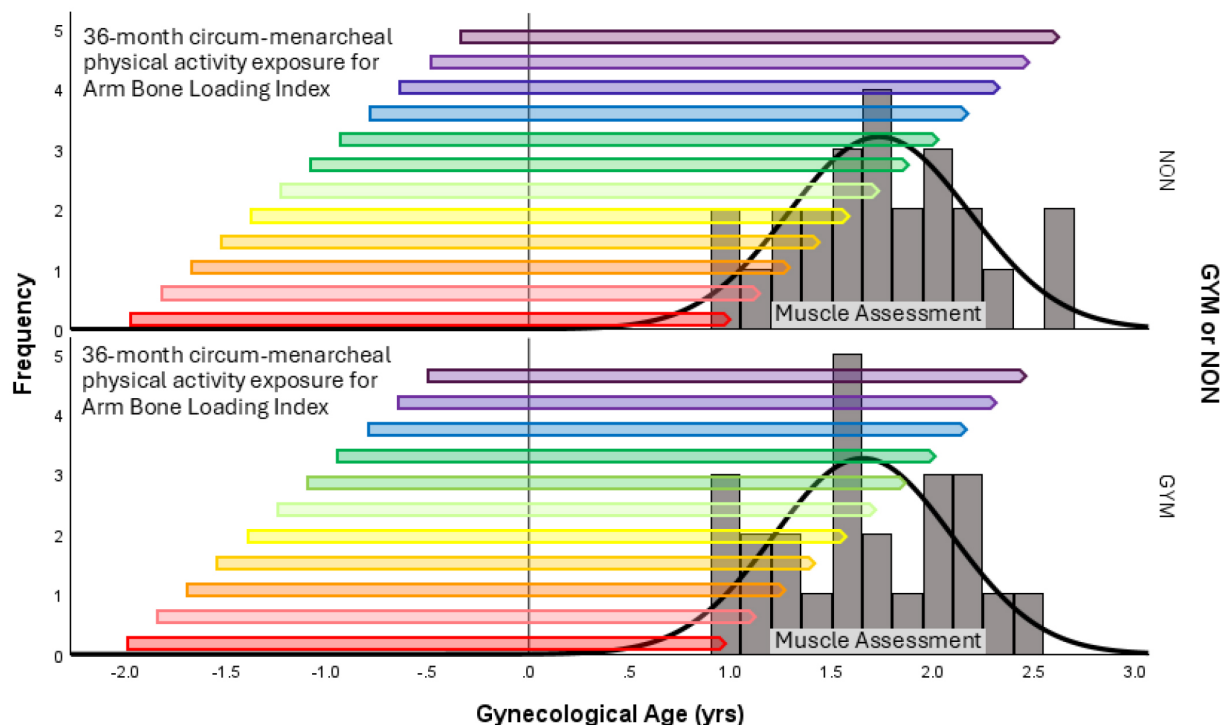


Figure 1. Data Collection Schematic. This data collection schematic depicts the 36-month period of organized physical activity exposure encompassed in the non-dominant Arm Bone Loading Index ($ARMBLI$). Exposures are shown as a spectrum of pre-menarcheal to post-menarcheal exposures (rainbow arrows), centered around the circum-menarcheal period of peak lean mass accrual velocity; variability is accounted for by incorporating gynecological age at muscle assessment in multiple linear regression models. Each participant's exposure period directly preceded their post-menarcheal session for assessment of non-dominant arm muscle properties and body size metrics. Session timings are roughly normally distributed, shown using gray histograms, paneled by group classifications as gymnasts (GYM: annual mean gymnastic training ≥ 6 h/wk for ≥ 1 year) and non-gymnasts (NON: all others). Each non-dominant arm muscle assessment session included contemporaneous measurements of height (cm), weight (kg), ulna length (mm), arm non-bone fat-free mass (kg, Hologic whole-body DXA), grip strength (kg, Takei dynamometer), and flexion and extension peak torques for elbow and shoulder (Nm, Biodex dynamometer).

Methods

Our longitudinal research protocols have been conducted since 1997, in compliance with the Declaration of Helsinki. From 2008 onward, we quantified muscle function using grip strength and isokinetic testing; both represent upper body strength with demonstrated clinical value and high reliability^{12,13,39}. Longitudinal participants included in the present analysis comprised “Cohort B” from our validation study (recruited 2008-2013)³⁸, who contributed: 1) circum-menarcheal activity data for $ARMBLI$; 2) post-menarcheal measures of anthropometry and non-dominant arm grip strength (kg), peak torques (Nm), and DXA $ARMNBFFM$ (kg), measured contemporaneously, roughly 1-2.5 years post-menarche (See Figure 1- Schematic). “Cohort A” girls (recruited 1997-2003) were excluded from the present analysis due to lack of required muscle functional data.

We published a validation study of $ARMBLI$, representing exposures to over 50 organized physical activities; for each girl, this metric quantifies osteogenic exercise for the non-dominant arm based on the sum of all types and cumulative hours of physical activity reported across the 3-year circum-menarcheal period³⁸. The non-dominant arm emphasizes musculoskeletal adaptation to organized physical activities rather than activities of daily living, amplifying effect sizes for $ARMBLI$ associations with muscle properties across the cohort. Because of the unique loading imparted by gymnastic training (near-symmetric loading of both arms with high weight-bearing impacts), we also evaluated “gymnasts” and “non-gymnasts” as separate sub-samples. Girls exposed to ≥ 12 months of artistic gymnastics, annual mean training ≥ 6 hours per week, were categorized as gymnasts (GYM). Others, who engaged in a broad variety of physical activities at varying doses³⁸, were

categorized as non-gymnasts (NON). Accordingly, some NON participated in gymnastics, and most GYM participated in activities other than gymnastics.

We calculated ARM_{BLI} based on the sum of exposure hours for all organized physical activities reported using semi-annual calendar-based questionnaires for the 36-month circum-menarcheal period³⁸. Each activity was rated on magnitude, frequency, velocity, and involvement of the non-dominant arm by a multidisciplinary panel of professionals and advanced trainees³⁸. Magnitude (score 0-5) reflects the mass of loads applied to the non-dominant arm, as well as the act of gripping, and/or use of a lever (e.g. bats, sticks, racquets). Velocity (score 0-5) represents the rate at which forces are conferred during load application (impact: zero, low, medium, high). Frequency (score 1-6) reflects how often loads are applied during training. To account for disproportionate dominant arm loading, a “non-dominance” value was applied [0.33= primarily dominant arm-loading (e.g. badminton, frisbee); 0.66= moderate non-dominant arm loading (e.g. basketball, hockey); 1.0= near-symmetric bilateral loading (e.g. gymnastics, powerlifting, soccer)]³⁸. Cumulative training exposure was based on semi-annual physical activity records, detailing hours per week of activity-specific exercise participation, month-by-month, for 36 months prior to the focal study visit. Cumulative ARM_{BLI} was calculated as the 3-year sum of activity-specific BLI values across all activities recorded for each participant: $ARM_{BLI} = \sum [(Magnitude + Velocity) \times Frequency \times Training Exposure \times Non-dominance]$. Additional details, including activity-specific loading ratings and Cohort B physical activity descriptive statistics are reported in our validation study and appendices³⁸.

For the present study, included participants attended muscle functional assessment visits roughly +1.0 to +2.5 years post-menarche (See Figure 1 – Schematic). Menarche dates had been documented during semi-annual prospective queries to calculate gynecological age at muscle assessment (GYN_{AGE} : yrs, centered at menarche = age at assessment minus age at menarche). At the muscle assessment, height (cm, wall-mounted stadiometer), non-dominant ulna length (mm, ruler from ulnar styloid to olecranon, rounded to nearest 5mm), and weight (kg, digital scale) were also recorded. Whole-body DXA scans included the non-dominant arm region of interest (Hologic Discovery A, Waltham MA; humeral head separated from scapula/trunk at glenoid fossa) to calculate non-dominant arm non-bone fat-free mass [arm lean mass – bone mineral content = ARM_{NBFFM} , kg; root mean square CV 2.0%]. Non-dominant arm grip strength was measured using a handheld dynamometer (GRIP: kg, best of 3 attempts: Takei Model 5001).

Non-dominant arm peak torques (PKTQ, Nm) were measured using standard isokinetic concentric/concentric dynamometer procedures to assess shoulder flexion (SF), shoulder extension (SE), elbow flexion (EF), and elbow extension (EE) muscle force generation (Biodex Medical Systems, Shirley, NY: pediatric attachments, as needed,

See Supplementary Table 1 for details). For each PKTQ measure, two sets of three repetitions were performed, yielding 6 repetitions per outcome per participant [shoulder = $60^{\circ}\cdot s^{-1}$; elbow = $30^{\circ}\cdot s^{-1}$]. If the coefficient of variation (CV) for repetitions within a given set exceeded 15% (shoulder) or 20% (elbow), an extra set was performed; if 3 sets were performed without valid CVs, data were excluded for that PKTQ outcome. Peak torque data were filtered, windowed, and gravity-corrected to yield the highest valid results as dependent variables (PKTQ: SF, SE, EF, EE). A single investigator (CS), trained in appropriate Biodex testing methods, ensured that standardized procedures were applied, including calibration, warm-up, positioning, stabilization, range of motion limits, axis alignment, rest intervals, and participant encouragement, yielding mean coefficients of variation <10% [EF 8.3%, EE 9.8%, SF 7.8%, SE 9.4%].

We used IBM SPSS (v31.0.0.0_117) to run the power analysis, screen for normality of data distributions, generate descriptive statistics, and perform correlations and regressions. Mann-Whitney U tests assessed group distribution differences. A Pearson correlation matrix was generated (one-tailed). We hypothesized positive correlations for body size/hypertrophy metrics and ARM_{BLI} vs. muscle function outcomes; we purposefully constrained the gynecological age range to limit physical maturity variance and influence. Multiple linear regression analyses evaluated associations of independent variables with dependent variables for the main hypothesis tests (GRIP, PKTQ). Multiple linear regressions were run using 2 main model formats to assess the independent explanatory value of ARM_{BLI} as a predictor of post-menarcheal GRIP and PKTQ: 1) Body Size Models (Model 1) accounted for variation in physical maturity (menarche-centered, gynecological age: GYN_{AGE} , yrs) and body size (height, weight); 2) Local Hypertrophy Models (Model 2) accounted for the influence of physical maturity, lever length and local muscle hypertrophy (GYN_{AGE} , ulna length, ARM_{NBFFM}). Each model was run for the TOTAL analysis sample, GYM subsample (gymnast classification), and NON subsample (non-gymnast classification). Regression results are presented as unstandardized coefficients (B) with 95% confidence intervals in Table 3; standardized coefficients (St B) are provided in Supplementary Tables 2–6. Semi-partial correlation coefficients (SPCC) and squared SPCC (R^2) indicate the proportion of independent variance explained by independent variables. Model assumptions were checked to ensure that regression model residuals were roughly normally distributed, and Cook’s distances were <0.5 (2 exceptions <0.59). All variance inflation factors were <5, indicating acceptable multicollinearity. For this exploratory analysis, power analysis indicated required $n=24$, based on the effect size observed for ARM_{NBFFM} in a model including GYN_{AGE} , Height, Weight, and ARM_{BLI} [partial $r=0.525$; power 0.818; $\alpha=0.05$]. Based on these estimates, our TOTAL sample analyses of the primary outcome (GRIP) and secondary outcomes (PKTQs) should be adequately

Table 1. Descriptive Statistics for GYM vs. NON comparisons.

Independent Variables													
VARIABLE NAME	TOTAL (n=48)				GYM (n=24)				NON (n=24)				MWU p
	Min	Max	Mean	sd	Min	Max	Mean	sd	Min	Max	Mean	sd	
Chronological Age (yrs)	11.3	17.2	14.4	1.2	12.2	17.2	14.7	1.0	11.3	16.5	14.0	1.4	0.083
Gynecological Age (yrs)	0.9	2.6	1.7	0.4	1.0	2.4	1.65	0.4	0.9	2.6	1.7	0.4	0.543
Height (cm)	144.2	174.5	161.3	6.7	144.2	170.5	160.0	6.5	152.6	174.5	162.7	6.7	0.490
Weight (kg)	41.8	89.2	58.0	9.3	41.8	74.2	57.0	7.7	45.0	89.2	59.0	10.8	0.650
Ulna Length (mm)	230	280	251	13	230	270	249	12	240	280	253	13	0.454
Non-dominant Arm Non-bone FFM (kg)	1.5	3.7	2.2	0.4	1.7	3.7	2.3	0.4	1.5	3.0	2.0	0.4	0.035*
Annual Mean Non-aquatic Physical Activity (h/wk)	1.3	22.9	9.2	4.6	6.9	22.9	11.8	4.1	1.3	16.4	6.6	3.7	<0.001***
3-year Non-dominant Arm Bone Loading Index	0.6	32.5	9.1	7.7	2.6	32.5	15.0	6.7	0.6	10.6	3.3	2.4	<0.001***
Dependent Variables													
VARIABLE NAME	TOTAL (n=48)				GYM (n=24)				NON (n=24)				MWU p
	Min	Max	Mean	sd	Min	Max	Mean	sd	Min	Max	Mean	sd	
Non-Dominant Hand Grip Strength (kg)	15.0	39.0	25.2	4.7	15.0	30.0	23.9	3.8	17.0	39.0	26.6	5.2	0.070
Elbow Extension Peak Torque (Nm)	6.3	45.1	24.5	8.5	6.3	45.1	25.3	9.8	14.5	39.4	24.6	6.5	0.509
Elbow Flexion Peak Torque (Nm) a,b,c	12.9	42.9	25.7	6.9	15.5	42.9	25.8	7.4	12.9	37.0	26.0	6.3	0.509
Shoulder Extension Peak Torque (Nm) d,e,f	2.8	86.6	33.6	14.9	2.8	86.6	36.1	19.0	12.3	53.0	31.2	9.2	0.358
Shoulder Flexion Peak Torque (Nm) g,h,i	16.8	56.8	28.5	7.9	16.8	56.8	30.4	9.2	17.6	40.2	26.7	6.0	0.124

*Gynecological age (yrs, centered at menarche); Ulna length (mm: recorded to nearest 5 mm, descriptive statistics are rounded to the nearest 1 mm). FFM= DXA-measured non-bone fat-free mass (Lean Mass – Bone Mineral Content). GYM = Gymnast group; NON = Non-gymnast group; EFPKTQ [a: TOTAL n= 47, b: GYM n=24, c: NON n=23]; SE PKTQ [d: TOTAL n= 45, e: GYM n=22, f: NON n=23]; SFPKTQ [g: TOTAL n= 46, h: GYM n=22 i: NON n=24]; sd = standard deviation; MWU = Mann-Whitney U test; *p < 0.05; ** p < 0.01; *** p < 0.001.*

powered; however, power for GYM & NON subsample analyses may be constrained, especially for secondary outcomes (PKTQs). Given multiple secondary outcomes and subgroup analyses, p-values were not adjusted for multiplicity and should be interpreted as exploratory.

Results

Descriptive Statistics and Mann-Whitney U Tests

All 48 Cohort B participants met inclusion criteria, providing data for our primary outcome, non-dominant hand grip strength (TOTAL n=48, GYM n=24, NON n=24, adequate sample sizes based on power estimates). During PKTQ measurements, GYM and NON provided complete data for EE, but EF data were reduced by one high CV (TOTAL n= 47; NON n= 23, GYM n=24). SE data were reduced by 2 CV exclusions and 1 gymnast injury (TOTAL n=45: NON n=23; GYM n=22). SF data were reduced by the same gymnast injury and 1 CV exclusion (TOTAL n=46; NON n=24; GYM n= 22). GYM and NON CV distributions were comparable for all PKTQ variables, before and after exclusion of individuals with CVs outside the acceptable range (Mann-Whitney U, p>0.3). Descriptive statistics are shown in Table 1 with Mann-

Whitney U results for GYM vs. NON group comparisons. The mean age at muscle functional assessment was 14.4 years. All participants reported menarche before age 16 years (no primary amenorrhea). GYM had higher $ARMNB_{FFM}$ and $ARMBLI$ than NON (Mann-Whitney U, p<0.04). We detected no significant differences for GYM vs. NON distributions of $GYNAGE$, height, weight, ulna length, or muscle function metrics (Mann-Whitney U, p>0.12), but we detected a non-significant bias towards higher age and lower GRIP in GYM vs. NON (Mann-Whitney U, p<0.09).

Associations of Key Predictors

One-tailed Pearson correlation results (Table 2) show strong positive associations among height, weight, and ulna length. Similarly, $ARMNB_{FFM}$ was significantly, positively correlated with height, weight, and ulna length ($r=+0.37$ to $+0.66$, p<0.01), as well as GRIP, EF, EE, SF, and SE ($r=+0.32$ to $+0.73$, p<0.001 to p<0.05). In contrast, $ARMBLI$ only correlated with $ARMNB_{FFM}$ ($r= +0.347$, p=0.008). $ARMBLI$ showed non-significant correlations with SE and SF ($r=+0.18$, $+0.22$, p=0.12 to 0.07) and was not correlated with EE and EF ($r < +0.08$, p>0.3). Surprisingly, $ARMBLI$ showed a weak negative correlation with GRIP ($r= -0.27$, p=

Table 2. Summary of Pearson Correlation Results (r).

	ARMBLI n= 48	GYNAGE (yrs) n= 48	Height (cm) n= 48	Weight (kg) n= 48	Ulna Length (mm) n= 48	ARMNBFFM (kg) n= 48	GRIP (kg) n= 48	EF (Nm) n= 47	EE (Nm) n= 48	SF (Nm) n= 46	SE (Nm) n= 45
ARMBLI	1	-0.096	-0.229	-0.052	-0.139	0.347**	-0.272*	0.061	0.073	0.225	0.178
GYNAGE (yrs)		1	0.142	0.091	0.041	0.137	0.175	0.103	0.144	0.174	0.125
Height (cm)			1	0.594***	0.838***	0.367**	0.457***	0.339**	0.177	0.268*	0.224
Weight (kg)				1	0.605***	0.659***	0.540***	0.511***	0.360**	0.559***	0.455***
Ulna length (mm)					1	0.477***	0.445***	0.400**	0.265*	0.257*	0.197
ARMNBFFM (kg)						1	0.317*	0.577***	0.438***	0.730***	0.538***

*ARMBLI = Non-dominant arm bone loading index for 3yrs circum-menarche; GYNAGE= gynecological age (yrs, centered at menarche). ARMNBFFM = DXA-measured non-dominant arm non-bone fat-free mass (kg: Lean Mass – Bone Mineral Content). GRIP = Non-dominant hand grip strength (kg); Non-dominant arm Biodex peak torques (Nm): EF (elbow flexion); EE (elbow extension); SF (shoulder flexion); SE (shoulder extension). One-tailed significance results, hypothesizing positive correlations, **negative correlation presented in bold italics**; *p < 0.05; ** p < 0.01; *** p < 0.001.*

Table 3. Summary of Non-dominant Arm Bone Loading Index Predictive Value from Multiple Linear Regression: Unstandardized B values [95%CI] with % of independent variance explained (squared semi-partial correlation coefficients).

MODEL 1: Body Size Model						
	Sample	Grip Strength (kg, n=48)	EF Peak Torque (Nm, n=47) a,e	EE Peak Torque (Nm, n=48)	SF Peak Torque (Nm, n=46) b,d	SE Peak Torque (Nm, n=45) c,d,e
MODEL 1	TOTAL (n=48) a,b,c	-0.127 [-0.281, +0.026] 4.1%	+0.099 [-0.147, +0.345] 1.1%	+0.104 [-0.219, +0.426] 0.8%	+0.270* [+0.018, +0.523] 6.8%	+0.399 [-0.137, +0.934] 4.2%
	GYM (n=24) d	-0.04 [-0.317, +0.237] 0.4%	+0.188 [-0.256, +0.631] 2.4%	+0.273 [-0.359, +0.906] 2.7%	+0.020 [-0.445, +0.486] 0.00%	+0.043 [-0.882, +0.967] 0.00%
	NON (n=24) e	-1.127** [-1.868, -0.387] 17.7%	-0.371 [-1.717, +0.976] 1.3%	-0.388 [-1.946, +1.170] 1.3%	-0.824 [-1.840, +0.193] 7.1%	+0.572 [-2.003, +3.146] 1.0%
MODEL 2: Local Hypertrophy Model						
	Sample	Grip Strength (kg, n=48)	EF Peak Torque (Nm, n=47) a,e	EE Peak Torque (Nm, n=48)	SF Peak Torque (Nm, n=46) b,d	SE Peak Torque (Nm, n=45) c,d,e
MODEL 2	TOTAL (n=48) a,b,c	-0.202* [-0.383, -0.021] 8.1%	-0.092 [-0.348, +0.165] 0.8%	-0.066 [-0.417, +0.286] 0.3%	-0.072 [-0.320, +0.175] 0.4%	-0.085 [-0.678, +0.508] 0.1%
	GYM (n=24) d	-0.041 [-0.327, +0.245] 0.4%	+0.022 [-0.413, +0.457] 0.0%	+0.155 [-0.497, +0.806] 0.8%	-0.212 [-0.652, +0.227] 2.1%	-0.282 [-1.454, +0.890] 0.9%
	NON (n=24) e	-0.935* [-1.772, -0.099] 11.6%	-0.400 [-1.775, +0.976] 1.4%	-0.466 [-2.040, +1.107] 1.8%	-0.392 [-1.586, +0.803] 1.5%	+0.861 [-1.768, +3.490] 2.4%

*Model 1 = Body Size Model [gynecological age (yrs, centered at menarche), height (cm), weight (kg), ARMBLI].
Model 2 = Local Hypertrophy Model [gynecological age (yrs), ulna length (mm), arm non-bone fat-free mass (kg), ARMBLI].
* p < 0.05; ** p < 0.01; *** p < 0.001;
GRIP = Grip Strength (kg); Biodex Peak torques (Nm): EF = Elbow Flexion; EE = Elbow Extension; SF = Shoulder Flexion; SE = Shoulder Extension.
a: TOTAL n=47 for EF, b: TOTAL n=46 for SF, c: TOTAL n=45 for SE, d: GYM n=22 for SF & SE; e: NON n=23 for EF & SE.*

0.031: opposite direction to the one-tailed test hypothesis).

To specifically evaluate the influence of local muscle hypertrophy for each functional variable without ARM_{BLI} , we entered GYN_{AGE} , ulna length, and ARM_{NBFFM} into basic regression models. Surprisingly, ARM_{NBFFM} did not exhibit independent explanatory value for GRIP (SPCC= +0.10, $p=0.46$), as GRIP was dominated by ulna length (SPCC= +0.34, $p<0.02$). In contrast, all PKTQs were dominated by ARM_{NBFFM} (SPCC: EE= +0.34, EF= +0.43, SE= +0.50, SF= +0.67; $p<0.02$).

Multiple Linear Regression Model Analyses

Supplementary Tables 2-6 display regression output for GRIP, EF, EE, SF, and SE, presenting different variable combinations, as Model 1 (Left: Body Size) and Model 2 (Right: Local Hypertrophy) for the full cohort (TOTAL: top) and subgroups (GYM: middle; NON: bottom). Table 3 summarizes ARM_{BLI} unstandardized B coefficients with 95%CI and squared semi-partial correlation coefficients across models and variables.

Supplementary Table 2 details results for the primary outcome of grip strength. Left panels assess GRIP as a function of ARM_{BLI} , accounting for whole-body size (Model 1: GYN_{AGE} , height, weight, ARM_{BLI}). In TOTAL Model 1, only weight was a significant, independent, positive predictor of GRIP (SPCC= +0.35, $p=0.005$), with a non-significant, negative association for ARM_{BLI} vs. GRIP (SPCC= -0.20, $p=0.101$). In GYM Model 1, we detected no significant, independent predictors of GRIP. In NON Model 1, significant, independent, positive predictors of GRIP included height (SPCC= +0.39, $p=0.008$) and weight (SPCC= +0.31, $p=0.032$), whereas ARM_{BLI} was a significant, independent, negative predictor of GRIP (SPCC= -0.42, $p=0.005$). Supplementary Table 2 (right) displays GRIP results, accounting for local hypertrophy (Model 2: GYN_{AGE} , ulna length, ARM_{NBFFM} , ARM_{BLI}). In TOTAL Model 2, ARM_{BLI} was a significant, independent negative predictor of GRIP (SPCC= -0.28, $p=0.029$), while ARM_{NBFFM} only approached a positive, independent GRIP association (SPCC= +0.23, $p=0.079$). In GYM Model 2 ($n=24$), no key variables were significant, independent GRIP predictors. In NON Model 2 ($n=24$), ARM_{BLI} was a significant independent, negative GRIP predictor (SPCC= -0.34, $p=0.030$).

Supplementary Tables 3 and 4 detail results for the secondary outcomes of elbow PKTQs. Left panels assess body size dynamics for EF and EE (Model 1: GYN_{AGE} , height, weight, ARM_{BLI}). In TOTAL Model 1, weight was the only significant, independent, positive predictor of EF (SPCC= +0.37, $p=0.008$) and EE (SPCC= +0.30, $p=0.036$). In GYM Model 1, weight remained a significant, positive, independent predictor of EF (SPCC= +0.51, $p=0.008$) and EE (SPCC= +0.49, $p=0.013$). In NON Model 1, no key variables were significant, independent predictors of EF or EE. Right panels account for hypertrophy in EF and EE (Model 2: GYN_{AGE} , ulna length, ARM_{NBFFM}). In TOTAL Model 2, ARM_{NBFFM} was the only significant, independent, positive

predictor of both EF and EE (respectively: SPCC= +0.42, +0.32; $p<0.03$). In GYM Model 2, ARM_{NBFFM} remained a significant, independent, positive predictor of both EF and EE (respectively: SPCC= +0.54, +0.42; $p<0.04$). In NON Model 2, no key variables were independent predictors of elbow PKTQs, but a non-significant, small, positive effect was observed for height versus EF (SPCC= +0.28, $p=0.183$). Thus, after accounting for body size or local hypertrophy, ARM_{BLI} was not an independent predictor of EF or EE in any sample.

Supplementary Tables 5 and 6 detail results for the secondary outcomes of shoulder PKTQs. Left panels assess body size dynamics (Model 1: GYN_{AGE} , height, weight, ARM_{BLI}). In TOTAL Model 1, weight was a significant, independent, positive predictor of SF and SE (respectively: SPCC= +0.47, +0.38; $p\leq0.008$). ARM_{BLI} was also a significant, independent, positive predictor of SF only (SPCC= +0.26, $p=0.036$), with a non-significant, small, positive effect for prediction of SE (SPCC= +0.20, $p=0.141$). In GYM Model 1, weight remained a significant, independent, positive predictor of SF and SE (both SPCC= +0.70, $p<0.001$). In NON Model 1, height was a significant, independent positive predictor of SF (SPCC= 0.40, $p=0.021$), and ARM_{BLI} showed a non-significant, small, negative effect for prediction of SF (SPCC= -0.27, $p=0.106$); no key variables were significant, independent predictors of SE in NON. Right panels account for hypertrophy in SF and SE (Model 2: GYN_{AGE} , ulna length, ARM_{NBFFM} , ARM_{BLI}). In TOTAL Model 2, ARM_{NBFFM} was the only significant, independent, positive predictor of SF and SE (respectively: SPCC= +0.62, +0.45; $p\leq0.002$). In GYM Model 2, ARM_{NBFFM} remained a significant, independent, positive predictor of SF and SE (respectively: SPCC= +0.74, +0.55; $p\leq0.008$). In NON Model 2, no key variables were significant, independent predictors of shoulder peak torques, but non-significant, positive, small effect sizes were seen for ARM_{NBFFM} versus SF (SPCC= +0.21, +0.25, $p<0.2$).

Table 3 shows that ARM_{BLI} was rarely a significant, independent predictor of muscle functional outcomes, serving as a negative predictor for the primary outcome of GRIP (TOTAL: only hypertrophy model; NON: body size & hypertrophy models) and a positive predictor of the secondary outcome of SF (TOTAL, only body size model), driven by a small effect in NON. Overall, ARM_{BLI} did not predict GYM muscle function indices.

Discussion

Despite our validation study showing that circum-menarcheal arm-specific osteogenic activity exposure was an independent, positive predictor of post-menarcheal arm muscle mass, this osteogenic activity index (ARM_{BLI}) was not a consistent, independent predictor of local muscle function at the same post-menarcheal time point. Thus, our main hypothesis was not supported in this exploratory analysis. In fact, ARM_{BLI} was a negative predictor of our primary outcome of grip strength in the full cohort, driven

mainly by non-gymnast patterns – a direct contradiction of our hypothesized positive correlation. In contrast, ARM_{BLI} was an independent, positive predictor of shoulder flexion peak torques in the full cohort, with a suggestion of positive predictive value for shoulder extension, wholly explained by local hypertrophy, and limited by constrained sample size. Surprisingly, ARM_{BLI} did not independently predict elbow peak torques in any model. In the isolated subsample of gymnasts, ARM_{BLI} demonstrated no independent predictive value for any muscle functional outcome.

In our prior work, ARM_{BLI} accounted for roughly 20% of variance in ARM_{NBFFM} , after accounting for gynecological age and height in the same Cohort B analysis sample³⁸. In the current analysis, ARM_{NBFFM} Pearson correlations indicated that local muscle hypertrophy was a significant factor in elbow and shoulder peak torques after menarche, but regression results were less straight-forward. For the TOTAL Cohort B sample, ARM_{NBFFM} explained 12% to 45% of variance in peak torques for both elbow and shoulder before adding ARM_{BLI} ; after accounting for ARM_{BLI} , ARM_{NBFFM} retained 10% to 38% of peak torque variance. Some arm muscles used in shoulder dynamometry lie fully within the DXA arm region of interest that assesses ARM_{NBFFM} (SF: coracobrachialis, biceps brachii), while other key shoulder girdle muscles extend only partially into the DXA arm region of interest [pectorals (SF), deltoids (SE/SF), teres major (SE), latissimus dorsi (SE)]. Thus, we hypothesize that ARM_{BLI} exhibits a small, positive effect size for independent prediction of SF in the TOTAL sample, supplementing ARM_{NBFFM} , because the latter does not fully encapsulate pectoral and deltoid muscle hypertrophy. In contrast, SE variance appears to be fully explained by ARM_{NBFFM} , despite only partial inclusion of relevant shoulder girdle musculature in the arm region of interest.

At the elbow, muscle mass of flexors (biceps brachii, brachioradialis, brachialis) and extensors (triceps brachii, anconeus) is fully encapsulated by ARM_{NBFFM} . Thus, for the full cohort, elbow joint muscle hypertrophy (mass) is a potent predictor of muscle function, and minimal variance remained to be explained independently by ARM_{BLI} . Notably, the explanatory value of arm muscle mass and whole-body weight associations was higher in the isolated gymnast subsample – a group that routinely supports their full body weight using upper extremity musculature during extreme maneuvers (GYM % variance explained for PKTQs: weight 24% 49%; ARM_{NBFFM} 18% to 54%). It makes sense that in gymnasts, osteogenic exercise loading would promote ARM_{NBFFM} , which encapsulates local hypertrophy well; thus, in GYM, ARM_{BLI} had no independent explanatory value for peak torques, after accounting for hypertrophy (ARM_{NBFFM}). Surprisingly, for non-gymnasts, neither local hypertrophy nor ARM_{BLI} independently explained muscle function for elbow peak torques, and grip strength patterns were counterintuitive.

Our thesis that the association between ARM_{BLI} and hypertrophy would extend to muscle functional advantages was not supported for grip strength in the full cohort.

Furthermore, grip strength was the only muscle functional outcome with a significant predictive model that held in the non-gymnast subsample, but contrary to our hypothesis, ARM_{BLI} was a negative predictor. Surprisingly, hypertrophy models exhibited lower explanatory value than body size models for grip strength; the significant influence of body weight in the full cohort appeared to be driven by weight and height associations in non-gymnasts alone. In a simple regression model for the full cohort, before adding ARM_{BLI} , grip strength was dominated by ulna length, with no independent predictive value of ARM_{NBFFM} ; after adding ARM_{BLI} to that simple hypertrophy model, ARM_{NBFFM} gained limited positive explanatory value for grip strength (ARM_{NBFFM} : 5% of variance explained, $p=0.08$). In the full cohort, ARM_{BLI} was negatively associated with grip strength, and non-gymnasts approached higher grip strength than gymnasts (Mann-Whitney U, $p=0.07$). In fact, for GYM alone, only 10-12% of grip strength variance was explained by either model, and no single independent variable explained significant grip variance. For grip strength assessment, we analyzed the highest of 3 trials as our dependent variable. It is possible that using the mean of 3 trials would have more directly assessed muscular stamina – a property that could amplify dose effects of chronic training.

We have several hypotheses for why grip strength patterns were comparatively unpredictable in the gymnast subsample. Uneven parallel bar training is the main “gripping” activity in artistic gymnastics; unless a gymnast was injured, bar work was not evaluated separately from training for other events or calisthenics work. Female gymnasts use a variety of types of “grip” devices, including “dowel grips”, on uneven parallel bars to boost gripping capacity, reduce fatigue, and improve maneuverability and control, while protecting against ripping injuries to skin⁴⁰. We did not document dowel grip use, but we hypothesize that variable “grip” device usage diminished positive associations between gymnastic training dose (cumulative hours) and observed grip strength in GYM. Furthermore, consistent or inconsistent use of dowel grips may have nullified positive predictive value of body weight and ARM_{NBFFM} for GRIP, in contrast to observed elbow and shoulder peak torque patterns. Future analyses should assess both maximum and mean grip strength to evaluate training benefits in young females and tease apart activities and training strategies that may boost or blunt grip-specific loading stresses and strains.

Inconsistent patterns for ARM_{BLI} vs. muscle function, especially GRIP, may also stem partly from the design of ARM_{BLI} to predict osteogenic loading, based on multiple parameters, including magnitude, frequency, velocity, and long-term exposure dose for each activity reported. Not all activities that boost ARM_{BLI} are dominated by “gripping” elements; for example, artistic gymnastics training and tumbling were assigned 40 arm bone loading units per training hour (ARM_{BLU}), the highest of any sport other than kickboxing (45 ARM_{BLU})³⁸. However, these activities have high bone loading due to bilateral impact loading, expected

to meet or exceed body weight forces, rather than promote grip strength: tumbling does not include any gripping; kickboxing involves “fist-making” contractions to support the hand and wrist for punching impacts rather than train GRIP directly; and as noted above, most female gymnastic GRIP training is via uneven bars, potentially limited by use of dowel grips. In contrast, many activities that more directly stimulate “GRIP” were assigned moderate bone loading units due to low impact (velocity) as for “heavy” upper extremity weight-training (30 ARMBLU), and/or lower weight-bearing (magnitude) in archery (20 ARMBLU) and lacrosse (18.5 ARMBLU)³⁸. In addition, mean reported training doses (cumulative hours per person) will have compounded ARMBLI differentials over the 3-year period, providing further contrast for gymnastics (1260 hrs), lacrosse (230 hrs), heavy weight-training (48 hrs), and archery (16 hrs)³⁸. These tactics appeared appropriate for explaining development of bone properties and ARMNBFFM, but were less effective for explaining muscle function, especially GRIP.

Limited ARMNBFFM explanatory value for muscle function may result from variable contributions of individual muscles and muscle groups as a proportion of local mass, as well as functional overlap of different muscles lumped together within this metric of hypertrophy. There may be substantial variability in the masses of specific muscles within the overarching ARMNBFFM measure. For example, three individuals may have identical ARMNBFFM, but girl X has small forearm muscles (low GRIP), girl Y has small triceps/biceps (low elbow PKTQs), and girl Z has small deltoids/pectorals (low shoulder PKTQs). Even if hypertrophy had high explanatory value for function within each muscle group, regression to the mean across the single, pooled index (ARMNBFFM) would mask underlying functional group explanatory value. However, Raymond-Pope’s 2020 analysis of male and female athletes (likely young adult, mean age 20 yrs) reported comparable explanatory value of DXA leg non-bone fat-free mass (LEGNBFFM) for Biodex knee flexion and extension PKTQs (KF, KE), regardless of whether LEGNBFFM was assessed as whole, upper, anterior, or posterior leg regions; at 60°·s⁻¹ settings, right leg correlations were generally highest for total LEGNBFFM (KF $r=+0.799$; KE $r=+0.591$), rather than more specific LEGNBFFM metrics⁴¹. Raymond-Pope’s results suggest that whole limb NBFFM explanatory value may be greater than regional metrics specific to the muscle functional group, even when predicting flexion/extension peak torques across involved joints. It is possible that similar findings apply to the elbow, where we observed large effect sizes for Pearson correlation coefficients with ARMNBFFM by Cohen’s standards (EF $r=+0.58$; EE $r=+0.44$); however, despite excluding portions of key shoulder musculature, our shoulder flexion correlations with ARMNBFFM were even larger, (SF $=+0.73$), while extension PKTQs were intermediate (SE $=+0.54$).

By Cohen’s standards, grip strength exhibited only a medium effect size for Pearson correlation with ARMNBFFM ($r=+0.32$). It is possible that grip strength would correlate

more directly with a more distal hypertrophy metric, such as forearm anatomical muscle CSA. Muscle CSA at the 65% site should specifically assess hypertrophy of relevant proximal forearm muscle bellies (flexor carpal radialis, flexor carpal ulnaris, flexor digitorum superficialis, and flexor digitorum profundus). Although 65% CSA excludes some hand muscles that are included in ARMNBFFM (flexor pollicis longus, flexor palmaris), those muscles do not contribute to Takei dynamometer grip strength. In a different sample (subset of Cohort A, $n=33/55$), we observed high post-menarcheal correlations for 65% muscle CSA vs. ARMNBFFM ($\rho=+0.62$, $p<0.001$), but weaker predictive value for muscle CSA than ARMNBFFM in “functional muscle-bone unit” analyses for the non-dominant arm⁴². In that prior analysis, after accounting for gymnastics history, gynecological age (2.5 to 4.5 yrs post-menarche), and height, ARMNBFFM was a significant, independent, positive predictor of 4% radius axial compressive strength index (29% of variance), arm BMC (41% of variance), and 33% radius bone geometry and theoretical strength indices (39%-50% of variance)⁴². In contrast, muscle CSA only predicted 4% radius axial compressive strength (38% of variance)⁴². As the 4% site is nearest to the hand/wrist, it is possible that, compared to ARMNBFFM, both muscle CSA and grip strength would be more closely tied to each other and 4% ultradistal bone strength. Unfortunately, we do not have complete, contemporaneous muscle CSA data for Cohort B to address this question directly in the current analysis.

Both ARMNBFFM and muscle CSA quantify hypertrophy, and they may share common genetic factors. However, they do not encapsulate other muscle functional determinants, such as muscle fiber type proportions, which may be partially genetically determined, or neuromuscular recruitment and firing^{4,43}. To address the latter in a small cross-sectional study of mixed maturity boys and girls, Gillen et al. used simultaneous electromyography, mechanomyography, and Biodex PKTQs to show that vastus lateralis neuromuscular adaptations explained additional variance in knee extension PKTQs beyond quadriceps muscle CSA⁴. In older male soccer players (mean age 20.6 years, unclear maturity), Masuda et al. showed that MRI-based local muscle CSA explained only 6% to 41% of variance in PKTQs for knee and hip flexion/extension, as well as hip abduction/adduction, leaving substantial muscle functional variance unexplained³. Both Gillen’s and Masuda’s results suggest that neuromuscular function may be important before peak muscle mass and CSA are achieved, but both studies were limited to lower extremity muscle function, with murky accounting for biological age, and limited or zero female representation.

After accounting for biological age and relevant size metrics in adolescent females using our full Cohort B sample, we see substantial unexplained variance for all muscle functional variables in our best regression models (GRIP: >62%; EE: >79%; EF: >64%; SE: >70%; SF: >44%). Even the most potent GYM model left 34% of variance in shoulder flexion unexplained, with 54% explained by ARMNBFFM hypertrophy. The most potent NON model was the body size

GRIP model – not even a function of muscle hypertrophy, as it excluded $ARMNB_{FFM}$. If neuromuscular adaptations routinely explain 30-40% of variance unexplained by $ARMNB_{FFM}$ hypertrophy, then one would expect substantial neuromuscular variance to be explained by activity dose – $ARMBLI$. However, independent $ARMBLI$ explanatory value was limited to the negative association of $ARMBLI$ vs. grip strength, explaining only 8% of variance in the full sample – an apparent detrimental effect of upper extremity-specific training dose representing >50 physical activities.

Surprisingly, we are not the first to observe a negative association between activity dose and grip strength. Burns et al. reported a negative association between upper extremity physical activity “dose” and GRIP in a sample of young males and females (mean age 10.5 years), contrasting with positive correlations between activity dose and a muscle function metric that incorporates both shoulder and elbow flexion⁴⁴. In their large sample of U.S. girls and boys from the 2012 National Youth Fitness Survey, Burns evaluated combined non-dominant + dominant handgrip strength and a modified pull-up test (supine row) relative to accelerometry-assessed moderate to vigorous physical activity dose (MVPA); their main predictors for fitness outcomes were peak 60 minute MVPA and average daily MVPA (1 to 7 days, non-dominant wrist accelerometry)⁴⁴. They reported a strong, negative correlation between GRIP and both peak MVPA and daily MVPA, dominated by 10th-90th percentile MVPA ranges; complex non-linear analyses showed non-significant MVPA-GRIP correlations within extreme MVPA ranges (bottom & top 10%). In contrast, MVPA metrics were positive predictors of modified pull-up results (repetitions in one minute), particularly among females, indicating potential dose-dependent benefits of arm-specific activity for muscle function in young girls of unclear puberty status⁴⁴. Thus, Burns’ results corroborate our negative associations of arm-specific activity dose with grip strength, contrasting with positive associations for SF and EF. Burns’ analyses were adjusted for BMI, so like ours, their results should not be a function of heavy body weight in less active youths. In both studies, children who were more active in the prior week may have been “fatigued” or dehydrated during GRIP testing, but it is unclear why these phenomena would not also adversely affect SF performance.

Another potential reason for inconsistent muscle function explanatory value may be the long $ARMBLI$ time course (3 years). This extended period is appropriate for densitometric assessment of bone adaptation, which requires >6 weeks to detect changes over multiple cycles of bone adaptive remodeling⁴⁵. Similarly, as lean mass accrual is rapid over this circum-menarcheal period, 3 years allowed detection of parallel loading adaptations in local skeletal muscle mass³⁸. In a mixed longitudinal analysis of 8 years of growth data centered at peak height velocity, Baxter-Jones et al. assessed regional NB_{FFM} gains based on validated semi-quantitative physical activity ratings (range PA1-PA5: recorded 2-3 times per year, 8 years)⁴⁶. After accounting for height and weight, no significant interaction

with biological age was seen; for each PA increment, females gained 31.4g $ARMNB_{FFM}$ (up to 157g gained for PA5) and 162.9g $LEGNB_{FFM}$ (up to 814.5g for PA5), while boys gained 120% more $ARMNB_{FFM}$ but only 21% more $LEGNB_{FFM}$ per PA increment⁴⁶. These findings suggest that sex-specific hypertrophic differences for arms & legs may not be fully captured by their general PA metric; unfortunately, they did not evaluate muscle functional data. In contrast, our 3-yr arm-specific loading metric was associated with incremental differences of 24 g $ARMNB_{FFM}$ per $ARMBLI$ unit (range of variance explained: 0 to 768g, after adjustment for gynecological age and height) in Cohort B, or 672g after accounting for gynecological age, height, and weight³⁸. Thus, we observed a broader incremental range for upper extremity PA dose gains, in line with Baxter-Jones’ female leg range and male arm range. Our broader arm range may be related to our arm-specific PA metric and the longer exposure period (3-yr dose vs. year-to-year level), as well as high representation of gymnasts whose arms function more like legs (extreme whole-body weight-bearing and impact-loading during inverted maneuvers). The fundamentally different loading patterns of weight-bearing and impact conferred during artistic gymnastics training may yield differences in bone, lean mass, and muscle function that are encompassed well by $ARMBLI$ in the context of the TOTAL cohort but provide limited discriminatory value within the GYM sub-group.

In contrast to bone and lean mass which accumulate during growth, muscle function may decline rapidly with detraining. Time courses for loss of muscular function in pre-pubertal and early pubertal boys and girls have been seen over detraining periods of 6-9 weeks^{47,48}. Accordingly, for girls with high arm-specific loading during the first 1-2.5 years, $ARMBLI$ could appear high, even if substantial detraining occurred over 6-12 months immediately prior to GRIP and PKTQ assessments, thereby reducing explanatory value of the 3-year osteogenic activity dose for post-menarcheal function. This phenomenon would be exaggerated for girls who participate in high loading activities such as gymnastics or kick-boxing but are injured or quit in the months immediately prior to the functional assessment.

Detraining effects for muscle function may vary by physical maturity. In a study of males and females (mainly pre-puberty), 9 weeks of shoulder resistance training (extension & flexion) saw training gains of 96% for SF and 33% for SE, with higher retained gains after 9 weeks of detraining for SF (66%) vs. SE (-0.4%)⁴⁸. Perhaps maintenance of SF training benefits across a longer time course explains why SF was the only functional metric with positive $ARMBLI$ associations, independent of body mass (Model 1), but largely a function of hypertrophy ($ARMNB_{FFM}$, Model 2). In contrast, across 12 months of upper extremity, dumbbell resistance training in young women, Heinonen et al. failed to detect benefits in grip strength and bone densitometric properties, but they detected significant EF PKTQ benefits that disappeared after 8 months of detraining⁴⁹. Future

analyses should evaluate activity exposures that may vary over shorter intervals, immediately preceding muscle functional assessments, to more directly assess training gains in muscle function, parallel with bone adaptation and muscle hypertrophy.

Limitations

ARMBLI is limited by reliance on retrospective self-report data, but strengths include semi-annual data collection via prospective study design and documented high correlations with practice logs from gymnastic coaches. Validity is further supported by substantial, significant predictive value of ARMBLI for ARMNBFFM and forearm bone outcomes in our earlier work³⁸. As noted above, for GRIP, a 3-trial mean might serve as a better indicator of training benefits than “best of 3”; in particular, the latter contrasts with our practice of excluding PKTQ observations with high variability, which specifically limits the influence of deviations. Unfortunately, the methodological stringency of our PKTQ analyses may have weakened statistical power to evaluate those secondary outcomes. Future work should boost sample sizes to counteract potential attrition due to injuries and/or inconsistency across trials. Even for the primary outcome of GRIP, our modest sample size likely restricted power for some GYM subsample analyses. By design, 50% of girls had a history of substantial gymnastics training, with GYM reporting ≥ 1 hour of organized activity per day over the 3-year period (on average). These patterns likely meet or exceed youth MVPA recommendations⁵⁰. In contrast, <50% of non-gymnasts met this standard, even if mandated school physical education were added (roughly 1.5 hr/wk). This sample composition expands the range of arm-specific loading exposures but may limit generalizability to the broadly “inactive” U.S. population. Unfortunately, GYM bars-specific training hours and associated dowel grip use were not documented to allow evaluation of these potential factors in GRIP development. Also, fatigue from chronic heavy training in active girls may have depressed functional outcomes during testing. In contrast to tight controls used for short-term functional studies, our longitudinal protocols may have allowed wider testing variation related to dehydration, dietary factors, injuries, and motivation. Finally, we did not account for genetic variation or assess neuromuscular recruitment. Future pediatric muscle function studies should recruit larger samples to account for biological age and potential genetic and neuromuscular factors, in protocols that aim to standardize nutrition, hydration, and rest to ensure optimal muscle functional performance during testing.

Conclusion

Although our validated circum-menarcheal bone-loading index quantified long-term osteogenic loading of the non-dominant arm and associated muscle hypertrophy (ARMNBFFM), ARMBLI was an inconsistent independent

predictor of upper extremity muscle functional outcomes in the current exploratory analysis. In fact, before and after accounting for local hypertrophy, ARMBLI exhibited small to medium effect sizes for negative, independent associations with the primary outcome of grip strength, driven primarily by non-gymnasts, directly contradicting our hypothesized positive correlation. For the secondary outcomes of shoulder and elbow PKTQs, we observed small effect sizes for positive prediction by ARMBLI, apparently wholly encapsulated by local hypertrophy and/or body size. Future research is needed with larger sample sizes, representing a broad range of training exposures, with more tightly controlled muscular functional assessment protocols, to develop a more potent index of loading exposure. Researchers need to elucidate key factors beyond body size and local hypertrophy to quantify modifiable factors in the development of upper extremity muscle function. Prospective studies are scarce in adolescent females; future work should assess female muscle function as a function of heterogeneous exposures to diverse activity types, accounting for the influence of genetic factors, biological age, nutritional elements, hormonal variation, muscle hypertrophy, and neuromuscular recruitment.

Ethics Approval

Protocol #239475, approved by the Institutional Review Board of SUNY Upstate Medical University.

Consent to Participate

Girls and their parents provided written, informed assent and consent, respectively, for participation in our longitudinal study.

Authors' contributions

Brian Josephson (data management, data analysis, data interpretation, manuscript drafting and critical revision, final approval for publication, agreement to be accountable for all aspects of the work); Timothy Chan (data management, data analysis, data interpretation, manuscript drafting and critical revision, final approval for publication, agreement to be accountable for all aspects of the work); Carol A. Sames (Conception, design, data acquisition, data analysis, data interpretation, manuscript drafting and critical revision, final approval for publication, agreement to be accountable for all aspects of the work); Tamara A. Scerpella (Conception, design, data interpretation, manuscript drafting and critical revision, final approval for publication, agreement to be accountable for all aspects of the work); Jodi N. Dowthwaite (Conception, design, data acquisition, data analysis, data interpretation, manuscript drafting and critical revision, final approval for publication, agreement to be accountable for all aspects of the work, particularly data analysis).

Funding

The current project was supported by ongoing funding from the University of Wisconsin-Madison, School of Medicine and Public Health and past funding from the National Institute of Arthritis, Musculoskeletal, and Skin Diseases (RO1 AR54145).

Acknowledgments

We are extremely grateful for the dedication of our longitudinal study participants, and the support of their parents. We also value

the tremendous efforts of our study coordinator, Tina Craig, and our longstanding DXA technologists, Cathy Riley and Eileen Burd. We would like to thank Nat Ordway, M.Eng. and Chris Neville D.P.T. for their assistance with Biodex data collection.

References

- Lexell J, Sjoström M, Nordlund AS, Taylor CC. Growth and development of human muscle: a quantitative morphological study of whole vastus lateralis from childhood to adult age. *Muscle Nerve*. Mar 1992;15(3):404-9.
- Stricker PR, Faigenbaum AD, McCambridge TM, Council On Sports M, Fitness. Resistance Training for Children and Adolescents. *Pediatrics*. Jun 2020;145(6), e20201011.
- Masuda K, Kikuhara N, Takahashi H, Yamanaka K. The relationship between muscle cross-sectional area and strength in various isokinetic movements among soccer players. *J Sports Sci*. Oct 2003;21(10):851-8.
- Gillen ZM, Shoemaker ME, McKay BD, Bohannon NA, Gibson SM, Cramer JT. Muscle strength, size, and neuromuscular function before and during adolescence. *Eur J Appl Physiol*. Jul 2019;119(7):1619-1632.
- Casamento-Moran A, Fleeman R, Chen YT, et al. Neuromuscular variability and spatial accuracy in children and older adults. *J Electromyogr Kinesiol*. Aug 2018;41:27-33.
- Daly RM, Stenevi-Lundgren S, Linden C, Karlsson MK. Muscle determinants of bone mass, geometry and strength in prepubertal girls. *Med Sci Sports Exerc*. Jun 2008;40(6):1135-41.
- U.S. Centers for Disease Control and Prevention. [Internet] Student Physical Education and Physical Activity. 2024. Washington, D.C.: Department of Health and Human Services (US). [updated 2024 Jun 25; cited 2026 Jan 13] <https://www.cdc.gov/physical-activity-education/about/index.html>
- American Heart Association. [Internet] NFL Play 60. 2026. Dallas: American Heart Association. [cited 2026 Jan 13] <https://www.heart.org/en/professional/educator/nfl-play-60>
- Office of Disease Prevention and Health Promotion. [Internet] Walk. Run. Dance. Play. What's your move? 2026. Washington, D.C.: Department of Health and Human Services (US). [updated 2026 Jan 13; cited 2026 Jan 13] <https://odphp.health.gov/moveyourway>
- Smith JJ, Eather N, Morgan PJ, Plotnikoff RC, Faigenbaum AD, Lubans DR. The health benefits of muscular fitness for children and adolescents: a systematic review and meta-analysis. *Sports Med*. Sep 2014;44(9):1209-23.
- García-Hermoso A, Ramírez-Campillo R, Izquierdo M. Is Muscular Fitness Associated with Future Health Benefits in Children and Adolescents? A Systematic Review and Meta-Analysis of Longitudinal Studies. *Sports Med*. Jul 2019;49(7):1079-1094.
- Bohannon RW. Grip Strength: An Indispensable Biomarker For Older Adults. *Clin Interv Aging*. 2019;14:1681-1691.
- Fraser BJ, Schmidt MD, Huynh QL, Dwyer T, Venn AJ, Magnussen CG. Tracking of muscular strength and power from youth to young adulthood: Longitudinal findings from the Childhood Determinants of Adult Health Study. *J Sci Med Sport*. Oct 2017;20(10):927-931.
- Martinez-Gomez D, Welk GJ, Puertollano MA, et al. Associations of physical activity with muscular fitness in adolescents. *Scand J Med Sci Sports*. Apr 2011;21(2):310-7.
- Lloyd RS, Faigenbaum AD, Stone MH, et al. Position statement on youth resistance training: the 2014 International Consensus. *Br J Sports Med*. Apr 2014;48(7):498-505.
- Munhoz da Silveira Campos R, Moreira CE, Leao da Silva P, et al. The effect of aerobic plus resistance training associated with a long-term interdisciplinary weight loss program on visceral fat and isokinetic parameters in adolescents with obesity. *J Sports Med Phys Fitness*. Jun 2020;60(6):855-863.
- Barbosa CCL, da Costa JC, Romanzini CLP, et al. Interrelationship between muscle fitness in childhood and bone mineral density in adulthood: mediation analysis of muscle fitness in adulthood. *BMC Public Health*. Apr 4 2023;23(1):648.
- Naimo MA, Gu JK. The Relationship between Resistance Training Frequency and Muscle Quality in Adolescents. *Int J Environ Res Public Health*. Jul 1 2022;19(13): 8099.
- Eustace SJ, Page RM, Greig M. Isokinetic strength differences between elite senior and youth female soccer players identifies training requirements. *Phys Ther Sport*. Sep 2019;39:45-51.
- Mascarin NC, de Lira CAB, Vancini RL, de Castro Pochini A, da Silva AC, Dos Santos Andrade M. Strength Training Using Elastic Bands: Improvement of Muscle Power and Throwing Performance in Young Female Handball Players. *J Sport Rehabil*. May 2017;26(3):245-252.
- Risberg MA, Steffen K, Nilstad A, et al. Normative Quadriceps and Hamstring Muscle Strength Values for Female, Healthy, Elite Handball and Football Players. *J Strength Cond Res*. Aug 2018;32(8):2314-2323.
- Hadzic V, Erculj F, Bracic M, Dervisevic E. Bilateral concentric and eccentric isokinetic strength evaluation of quadriceps and hamstrings in basketball players. *Coll Antropol*. Sep 2013;37(3):859-65.
- Sugimoto D, Myer GD, Bush HM, Hewett TE. Effects of compliance on trunk and hip integrative neuromuscular training on hip abductor strength in female athletes. *J Strength Cond Res*. May 2014;28(5):1187-94.
- de Rezende A, II, Sampaio LHF, Bittar AJ, da Silva Hamu TCD, Wyon MA, Formiga C. The Relationship between Vitamin D Levels, Injury and Muscle Function in Adolescent Dancers. *Int J Sports Med*. Jun

- 2020;41(6):360-364.
25. Coratella G, Longo S, Ce E, et al. Sex-Related Responses to Eccentric-Only Resistance Training in Knee-Extensors Muscle Strength and Architecture. *Res Q Exerc Sport*. Sep 2018;89(3):347-353.
 26. Ralston GW, Kilgore L, Wyatt FB, Baker JS. The Effect of Weekly Set Volume on Strength Gain: A Meta-Analysis. *Sports Med*. Dec 2017;47(12):2585-2601.
 27. Mueller J, Mueller S, Stoll J, Baur H, Mayer F. Trunk extensor and flexor strength capacity in healthy young elite athletes aged 11-15 years. *J Strength Cond Res*. May 2014;28(5):1328-34.
 28. Almeida-Neto PF, de Medeiros RCDSC, de Matos DG, Baxter-Jones ADG, Aida FJ, de Assis GG, Silva Dantas PM, Cabral BGAT. Lean mass and biological maturation as predictors of muscle power and strength performance in young athletes. *PLoS One*. 2021 Jul 12;16(7):e0254552.
 29. Lesinski M, Prieske O, Granacher U. Effects and dose-response relationships of resistance training on physical performance in youth athletes: a systematic review and meta-analysis. *Br J Sports Med*. Jul 2016;50(13):781-95.
 30. Pereira S, Santos C, Tani G, Freitas D, Garbeloto F, Guimarães E, Robinson LE, Baxter-Jones A, Katzmarzyk PT, Maia J. Physical fitness spurts in pre-adolescent boys and girls: Timing, intensity and sequencing. *J Sports Sci*. 2022 Mar;40(6):630-637.
 31. Reyes A, Chaves R, Vasconcelos O, Pereira S, Tani G, Stodden D, Hedeker D, Maia J, Baxter-Jones A. Modeling the Dynamics of Children's Musculoskeletal Fitness. *Int J Environ Res Public Health*. 2023 Feb 8;20(4):2938.
 32. Daly RM, Saxon L, Turner CH, Robling AG, Bass SL. The relationship between muscle size and bone geometry during growth and in response to exercise. *Bone*. Feb 2004;34(2):281-7.
 33. Michalopoulou M, Kambas A, Leontsini D, et al. Physical activity is associated with bone geometry of premenarcheal girls in a dose-dependent manner. *Metabolism*. Dec 2013;62(12):1811-8.
 34. Pietrobello A, Faith MS, Wang J, Brambilla P, Chiumello G, Heymsfield SB. Association of lean tissue and fat mass with bone mineral content in children and adolescents. *Obes Res*. Jan 2002;10(1):56-60.
 35. Gabel L, Macdonald HM, Nettlefold L, McKay HA. Physical Activity, Sedentary Time, and Bone Strength From Childhood to Early Adulthood: A Mixed Longitudinal HR-pQCT study. *J Bone Miner Res*. Jul 2017;32(7):1525-1536.
 36. Herrmann D, Buck C, Sioen I, et al. Impact of physical activity, sedentary behaviour and muscle strength on bone stiffness in 2-10-year-old children-cross-sectional results from the IDEFICS study. *Int J Behav Nutr Phys Act*. Sep 17 2015;12:112.
 37. McKay HA, Bailey DA, Mirwald RL, Davison KS, Faulkner RA. Peak bone mineral accrual and age at menarche in adolescent girls: a 6-year longitudinal study. *J Pediatr*. Nov 1998;133(5):682-7.
 38. Dowthwaite JN, Dunsmore KA, Gero NM, et al. Arm bone loading index predicts DXA musculoskeletal outcomes in two samples of post-menarcheal girls. *J Musculoskelet Neuronal Interact*. Dec 2015;15(4):358-71.
 39. Wilk KE, Arrigo CA, Davies GJ. Isokinetic Testing: Why it is More Important Today than Ever. *Int J Sports Phys Ther*. 2024;19(4):374-380.
 40. Kara E, Sagioglu I, Vurgun H, et al. The Risk Factors Associated with Grip Lock Injuries in Artistic Gymnasts: A Systematic Review. *Int J Environ Res Public Health*. Feb 17 2023;20(4), 3589.
 41. Raymond-Pope CJ, Dengel DR, Fitzgerald JS, Bosch TA. Association of Compartmental Leg Lean Mass Measured by Dual X-Ray Absorptiometry With Force Production. *J Strength Cond Res*. Jun 2020;34(6):1690-1699.
 42. Dowthwaite JN, Kanaley JA, Spadaro JA, Hickman RM, Scerpella TA. Muscle indices do not fully account for enhanced upper extremity bone mass and strength in gymnasts. *J Musculoskelet Neuronal Interact*. Jan-Mar 2009;9(1):2-14.
 43. Roth SM. Genetic aspects of skeletal muscle strength and mass with relevance to sarcopenia. *Bonekey Rep*. 2012;1:58.
 44. Burns RD, Byun W, Bai Y, de Castro Silveira JF, Reuter CP. Dose-response associations of Monitor-Independent Movement Summary with health-related fitness in youth. *Scand J Med Sci Sports*. Nov 2023;33(11):2286-2298.
 45. Varley I, Ward M, Thorpe C, et al. External training load is associated with adaptation in bone and body composition over the course of a season in elite male footballers. *Bone Rep*. Jun 2023;18:101643.
 46. Baxter-Jones AD, Eisenmann JC, Mirwald RL, Faulkner RA, Bailey DA. The influence of physical activity on lean mass accrual during adolescence: a longitudinal analysis. *J Appl Physiol* (1985). Aug 2008;105(2):734-41.
 47. Faigenbaum AD, Westcott2 WL, Micheli LJ, et al. The Effects of Strength Training and Detraining on Children. *The Journal of Strength & Conditioning Research*. 1996;10(2):109-114.
 48. Sewall L, Micheli LJ. Strength training for children. *J Pediatr Orthop*. Mar-Apr 1986;6(2):143-6.
 49. Heinonen A, Sievänen H, Kannus P, Oja P, Vuori I. Effects of unilateral strength training and detraining on bone mineral mass and estimated mechanical characteristics of the upper limb bones in young women. *J Bone Miner Res*. 1996 Apr;11(4):490-501.
 50. U.S. Department of Health and Human Services. [Internet]. Physical Activity Guidelines for Americans, 2nd edition. 2018. Washington, DC: U.S. Department of Health and Human Services [cited 2026 Jan 13]. https://odphp.health.gov/sites/default/files/2019-09/Physical_Activity_Guidelines_2nd_edition.pdf

Supplementary Table 1. Standardized Biodex Data Collection Methodology.

Shoulder Positioning	The dynamometer orientation was set to 0°, dynamometer tilt 0°, seat orientation 15°, and seatback tilt 70°-80° based on comfort and set to proper axis of alignment. The axis of rotation was the acromial process in the sagittal plane. Stabilization used shoulder, waist and thigh straps. Range of Motion limits were set at 180° and 30°.
Elbow Positioning	The dynamometer orientation was set to 30°, dynamometer tilt 0°, positioning chair orientation 0°, and seatback tilt 85°. The axis of rotation was centered at the trochlea and capitulum, bisecting the longitudinal axis of the shaft of the humerus. Stabilization used shoulder, waist and thigh straps. The handgrip was placed to allow rotation through the full range of motion for the wrist (wrist ROM limits: 0° and 160°).
Gravity-correction Methods	Gravity correction procedures were performed with each girl placed in the standardized starting position prior to testing for both shoulder and elbow; automatic correction was performed by the Biodex dynamometer and recorded by its software, adding to extension and subtracting from flexion to compensate for the effect of gravity and yield an unbiased torque value.
Warm-up & Familiarization	For each joint, girls completed 3 flexion/extension movements through the full range of motion with no resistance (1 rep) and submaximal resistance (2 reps); then each girl verbally confirmed that they were comfortable and ready to start. Each participant had performed these measurement series on at least one prior occasion during at least one earlier longitudinal study year.
Rest Intervals	Other than 30 seconds for rest between repetitions, there were no additional pauses for rest between practice sets or testing sets within-joint. There was a break of approximately 4-5 minutes between elbow and shoulder testing to change attachments, adjust positioning, strap-in the participant, and realign equipment, as needed.
Standardized Encouragement	Participants faced away from the computer screen during testing, so no visual feedback was available for their performance. All girls received the following encouragements during the test: "pull up as hard/fast as you can; pull down as hard/fast as you can; good job".
Observer Blinding	The single biodex tester did not classify participants based on gymnast status, and they were not given details of GYM/NON classification on data collection sheets. Although participants may have mentioned gymnastics participation during data collection, some NON also participated in gymnastics during the 3 years of data collection leading up to and including the post-menarcheal muscle function assessment. Thus, gymnastic training did not fully map onto GYM classification.
Reliability	Acceptable Coefficients of Variation (CV) were ≤20% for elbow peak torques, ≤15% for shoulder peak torques. If a participants' CV was too high, a 3 rd trial was attempted; if their CV was still too high, their data were excluded. Single Observer, Mean CVs for included data: EF 8.3%, EE 9.8%, SF 7.8%, SE 9.4%

Supplementary Table 2. Multiple Linear Regression Model Results for Prediction of Post-menarcheal Grip Strength (kg).

SAMPLE	Model 1: Body Size Model						Model 2: Local Hypertrophy Model					
	Independent Variables	B	St B	Semi-Partial r	p	Model R ²	Independent Variables	B	St B	Semi-Partial r	p	Model R ²
TOTAL (n=48)	GYNAGE (yrs)	1.020	0.096	0.095	0.437	0.373	GYNAGE (yrs)	0.980	0.092	0.090	0.482	0.314
	Height (cm)	0.093	0.133	0.103	0.398		Ulna Length (mm)	0.093	0.252	0.205	0.112	
	Weight (kg)	0.223	0.442	0.353	0.005		ARMNBFFM (kg)	3.491	0.300	0.227	0.079	
	ARMBLI	-0.127	-0.210	-0.202	0.101		ARMBLI	-0.202	-0.332	-0.285	0.029	
GYM (n=24)	GYNAGE (yrs)	1.633	0.190	0.185	0.399	0.124	GYNAGE (yrs)	1.777	0.206	0.191	0.392	0.097
	Height (cm)	0.027	0.046	0.036	0.868		Ulna Length (mm)	0.048	0.158	0.137	0.536	
	Weight (kg)	0.110	0.224	0.184	0.401		ARMNBFFM (kg)	0.571	0.063	0.052	0.814	
	ARMBLI	-0.040	-0.071	-0.065	0.765		ARMBLI	-0.041	-0.072	-0.065	0.770	
NON (n=24)	GYNAGE (yrs)	-1.351	-0.117	-0.106	0.433	0.669	GYNAGE (yrs)	0.491	0.043	0.041	0.778	0.599
	Height (cm)	0.478	0.620	0.388	0.008		Ulna Length (mm)	0.192	0.483	0.245	0.108	
	Weight (kg)	0.199	0.413	0.306	0.032		ARMNBFFM (kg)	5.925	0.414	0.240	0.115	
	ARMBLI	-1.127	-0.515	-0.421	0.005		ARMBLI	-0.935	-0.427	-0.340	0.030	

*St B = standardized B coefficient; GYNAGE = gynecological age (yrs, centered at menarche);
 ARMBLI = $\Sigma [(Magnitude + Velocity) \times Frequency \times Non-dominance \times 3\text{-year Circum-menarcheal Sport-specific Training Exposure}]$.
 ARMNBFFM = DXA-measured non-bone fat-free mass of non-dominant arm (Lean Mass – Bone Mineral Content).
 All models: Normally distributed residuals; Cook's Distances <0.58; TOTAL & GYM models: Variance Inflation Factors <1.9; NON VIFs <2.6.*

Supplementary Table 3.

Multiple Linear Regression Model Results for Prediction of Post-menarcheal Biodex Elbow Flexion Peak Torque (Nm).

SAMPLE	Model 1: Body Size Model					Model R ²	Model 2: Local Hypertrophy Model					Model R ²
	Independent Variable	B	St B	Semi-Partial r	p		Independent Variable	B	St B	Semi-Partial r	p	
TOTAL (n=47)	GYNAGE (yrs)	0.966	0.061	0.061	0.647	0.277	GYNAGE (yrs)	0.394	0.025	0.024	0.844	0.363
	Height (cm)	0.085	0.083	0.064	0.627		Ulna Length (mm)	0.064	0.120	0.097	0.434	
	Weight (kg)	0.337	0.461	0.366	0.008		ARMNBFFM (kg)	9.429	0.552	0.422	0.001	
	ARMBLI	0.099	0.112	0.107	0.420		ARMBLI	-0.092	-0.103	-0.089	0.474	
GYM (n=24)	GYNAGE (yrs)	1.703	0.100	0.097	0.581	0.428	GYNAGE (yrs)	-0.224	-0.013	-0.012	0.943	0.469
	Height (cm)	-0.034	-0.030	-0.023	0.895		Ulna Length (mm)	0.046	0.075	0.066	0.698	
	Weight (kg)	0.610	0.626	0.514	0.008		ARMNBFFM (kg)	11.759	0.657	0.540	0.004	
	ARMBLI	0.188	0.167	0.154	0.387		ARMBLI	0.022	0.020	0.018	0.917	
NON (n=23)	GYNAGE (yrs)	-2.270	-0.158	-0.144	0.481	0.279	GYNAGE (yrs)	-0.943	-0.065	-0.062	0.753	0.320
	Height (cm)	0.409	0.439	0.277	0.183		Ulna Length (mm)	0.244	0.514	0.250	0.214	
	Weight (kg)	0.115	0.200	0.148	0.469		ARMNBFFM (kg)	2.276	0.131	0.073	0.712	
	ARMBLI	-0.371	-0.141	-0.116	0.570		ARMBLI	-0.400	-0.153	-0.119	0.549	

St B = standardized B coefficient; GYNAGE = gynecological age (yrs, centered at menarche);
 ARMBLI = $\Sigma [(Magnitude + Velocity) \times Frequency \times Non\text{-}dominance \times 3\text{-}year\ Circum\text{-}menarcheal\ Sport\text{-}specific\ Training\ Exposure]$.
 ARMNBFFM = DXA-measured non-bone fat-free mass of non-dominant arm (kg: Lean Mass – Bone Mineral Content).
 All models: Normally distributed residuals; Cook's Distances <0.5; TOTAL & GYM Variance Inflation Factors <1.9; NON VIFs <4.3.

Supplementary Table 4.

Multiple Linear Regression Model Results for Prediction of Post-menarcheal Biodex Elbow Extension Peak Torque (Nm).

SAMPLE	Model 1: Body Size Model					Model R ²	Model 2: Local Hypertrophy Model					Model R ²
	Independent Variable	B	St B	Semi-Partial r	p		Independent Variable	B	St B	Semi-Partial r	p	
TOTAL (n=48)	GYNAGE (YRS)	2.406	0.125	0.123	0.384	0.154	GYNAGE (YRS)	1.510	0.078	0.076	0.577	0.206
	Height (cm)	-0.057	-0.045	-0.035	0.805		Ulna Length (mm)	0.034	0.051	0.042	0.761	
	Weight (kg)	0.347	0.380	0.304	0.036		ARMNBFFM (KG)	8.933	0.424	0.321	0.023	
	ARMBLI	0.104	0.094	0.091	0.520		ARMBLI	-0.066	-0.060	-0.051	0.708	
GYM (n=24)	GYNAGE (YRS)	2.282	0.098	0.095	0.604	0.377	GYNAGE (YRS)	0.671	0.029	0.027	0.886	0.363
	Height (cm)	-0.140	-0.089	-0.070	0.705		Ulna Length (mm)	0.110	0.133	0.116	0.532	
	Weight (kg)	0.800	0.600	0.493	0.013		ARMNBFFM (KG)	12.545	0.512	0.421	0.033	
	ARMBLI	0.273	0.178	0.164	0.377		ARMBLI	0.155	0.101	0.091	0.625	
NON (n=24)	GYNAGE (YRS)	-0.351	-0.024	-0.022	0.922	0.063	GYNAGE (YRS)	0.329	0.023	0.022	0.920	0.095
	Height (cm)	0.238	0.247	0.154	0.496		Ulna Length (mm)	0.158	0.317	0.161	0.471	
	Weight (kg)	0.041	0.068	0.051	0.822		ARMNBFFM (KG)	0.802	0.045	0.026	0.907	
	ARMBLI	-0.388	-0.142	-0.116	0.608		ARMBLI	-0.466	-0.170	-0.135	0.542	

St B = standardized B coefficient; gynAGE = gynecological age (yrs, centered at menarche);
 ARMBLI = $\Sigma [(Magnitude + Velocity) \times Frequency \times Training\ Exposure \times Non\text{-}dominance]$.
 ARMNBFFM = DXA-measured non-bone fat-free mass of non-dominant arm (kg: Lean Mass – Bone Mineral Content).
 All models: Normally distributed residuals; Cook's Distances <0.5; TOTAL & GYM Variance Inflation Factors <1.9; NON VIFs <3.9.

Supplementary Table 5.

Multiple Linear Regression Model Results for Prediction of Post-menarcheal Biodex Shoulder Flexion Peak Torque (Nm).

SAMPLE	Model 1: Body Size Model						Model 2: Local Hypertrophy Model					
	Independent Variable	B	St B	Semi-Partial r	p	Model R ²	Independent Variable	B	St B	Semi-Partial r	p	Model R ²
TOTAL (n=46)	GYNAGE (yrs)	2.568	0.147	0.145	0.235	0.404	GYNAGE (yrs)	1.038	0.060	0.058	0.582	0.551
	Height (cm)	-0.047	-0.041	-0.032	0.794		Ulna Length (mm)	-0.086	-0.141	-0.114	0.280	
	Weight (kg)	0.495	0.587	0.469	<0.001		ARMNBFFM (kg)	15.691	0.813	0.616	<0.001	
	ARMBLI	0.270	0.270	0.261	0.036		ARMBLI	-0.072	-0.072	-0.062	0.558	
GYM (n=22)	GYNAGE (yrs)	1.919	0.096	0.093	0.552	0.604	GYNAGE (yrs)	-0.855	-0.043	-0.039	0.786	0.656
	Height (cm)	-0.415	-0.305	-0.238	0.137		Ulna Length (mm)	-0.122	-0.170	-0.148	0.312	
	Weight (kg)	1.057	0.861	0.703	<0.001		ARMNBFFM (kg)	19.362	0.894	0.737	<0.001	
	ARMBLI	0.020	0.015	0.014	0.927		ARMBLI	-0.212	-0.159	-0.145	0.323	
NON (n=24)	GYNAGE (yrs)	-1.165	-0.087	-0.079	0.621	0.534	GYNAGE (yrs)	1.226	0.092	0.089	0.623	0.390
	Height (cm)	0.563	0.632	0.395	0.021		Ulna Length (mm)	0.123	0.268	0.136	0.458	
	Weight (kg)	0.160	0.287	0.213	0.190		ARMNBFFM (kg)	7.050	0.426	0.247	0.184	
	ARMBLI	-0.824	-0.325	-0.266	0.106		ARMBLI	-0.392	-0.155	-0.123	0.501	

St B = standardized B coefficient; GYNAGE = gynecological age (yrs, centered at menarche);
 ARMBLI = $\Sigma [(Magnitude + Velocity) \times Frequency \times Non\text{-}dominance \times 3\text{-}year\ Circum\text{-}menarcheal\ Sport\text{-}specific\ Training\ Exposure]$.
 ARMNBFFM = DXA-measured non-bone fat-free mass of non-dominant arm (kg; Lean Mass – Bone Mineral Content).
 All models: Normally distributed residuals; Cook's Distances <0.5; TOTAL & GYM Variance Inflation Factors <1.9, NON VIFs <3.9.

Supplementary Table 6.

Multiple Linear Regression Model Results for Prediction of Post-menarcheal Biodex Shoulder Extension Peak Torque (Nm).

SAMPLE	Model 1: Body Size Model						Model 2: Local Hypertrophy Model					
	Independent Variable	B	St B	Semi-Partial r	p	Model R ²	Independent Variable	B	St B	Semi-Partial r	p	Model R ²
TOTAL (n=45)	GYNAGE (yrs)	3.108	0.091	0.089	0.517	0.261	GYNAGE (yrs)	1.372	0.040	0.039	0.768	0.300
	Height (cm)	-0.085	-0.038	-0.028	0.835		Ulna Length (mm)	-0.142	-0.119	-0.094	0.482	
	Weight (kg)	0.760	0.484	0.380	0.008		ARMNBFFM (kg)	21.802	0.607	0.449	0.002	
	ARMBLI	0.399	0.213	0.204	0.141		ARMBLI	-0.085	-0.045	-0.038	0.773	
GYM (n=22)	GYNAGE (yrs)	0.939	0.021	0.020	0.892	0.633	GYNAGE (yrs)	-0.501	-0.011	-0.010	0.956	0.429
	Height (cm)	-0.520	-0.186	-0.145	0.336		Ulna Length (mm)	-0.006	-0.004	-0.004	0.984	
	Weight (kg)	2.185	0.866	0.697	<0.001		ARMNBFFM (kg)	29.410	0.678	0.553	0.008	
	ARMBLI	0.043	0.016	0.014	0.924		ARMBLI	-0.282	-0.104	-0.093	0.618	
NON (n=23)	GYNAGE (yrs)	-3.646	-0.176	-0.154	0.485	0.155	GYNAGE (yrs)	-1.058	-0.051	-0.049	0.829	0.103
	Height (cm)	0.505	0.354	0.220	0.322		Ulna Length (mm)	0.016	0.021	0.010	0.963	
	Weight (kg)	0.001	0.001	0.001	0.997		ARMNBFFM (kg)	4.900	0.197	0.101	0.657	
	ARMBLI	0.572	0.112	0.101	0.646		ARMBLI	0.861	0.169	0.154	0.500	

St B = standardized B coefficient; GYNAGE = gynecological age (yrs, centered at menarche);
 ARMBLI = $\Sigma [(Magnitude + Velocity) \times Frequency \times Non\text{-}dominance \times 3\text{-}year\ Circum\text{-}menarcheal\ Sport\text{-}specific\ Training\ Exposure]$.
 ARMNBFFM = DXA-measured non-bone fat-free mass of non-dominant arm (kg; Lean Mass – Bone Mineral Content).
 All models: Normally distributed residuals; Cook's Distances <0.59; TOTAL & GYM Variance Inflation Factors <1.9, NON VIF <4.1.