
José Afonso ¹, Pantelis T. Nikolaidis ², Patrícia Sousa ¹ and Isabel Mesquita ¹

¹ Centre for Research, Formation, Innovation, and Intervention in Sport. Faculty of Sport – University of Porto, Portugal; ² Department of Physical and Cultural Education – Hellenic Army Academy, Athens, Greece

Abstract
Periodization is a core concept in training. Recently, systematic reviews and meta-analyses have attempted to provide a comprehensive overview of the topic, but theoretical criticisms have arisen with regard to how such research has been conducted. The purpose of the study was to review comprehensively the conceptual and methodological issues surrounding empirical research on periodization in training with human subjects. A search was conducted late in February 2016 on Academic Search Complete, CINAHL Plus, Medline, PsycINFO, PubMed, Scielo, Scopus, SPORTDiscuss and Web of Science. Forty-two randomized or randomized controlled trials were retrieved. Problems emerged in three domains: (a) Conceptually, periodization and variation were applied differently in research, while no empirical research tested predictions concerning direction, timing or magnitude of the adaptations; (b) Study design: More than 95% of papers investigated the ‘physical’ factor (mainly strength). Research on long-term effects was absent (no study lasted more than nine months). Controlling for confounding factors such as nutrition, supplementation and medication was largely ignored; (c) Data analysis was biased as dispersion in responsiveness was ignored when discussing the findings. Overall, research on periodization fails to analyze the conceptual premises proposed by these approaches.

Key words: Periodized programs, randomized trials, research paradigms.

Introduction
Periodization is a core scientific concept of training theory and methodology, and is widely acclaimed as being beneficial in exercise prescription, both for performance and health purposes (Issurin, 2008; Naclerio et al., 2013). It consists of the “systematic planning and structuring of training variables throughout designated training timeframes aimed at maximizing performance gains and minimizing the potential for overtraining or decrements in performance” (Harries et al., 2015, p. 1113). As its definition implies, periodization requires training variation (Gamble, 2006), but extends well beyond that. Indeed, it aims at achieving peak performances in certain, pre-specified periods in time (Fleck, 2008; Turner, 2011), while also avoiding overtraining and reducing the risk of injury (Naclerio et al., 2013).

It might be argued that performance or learning environments can rarely be well predicted in advance (Davids, 2012). Learning settings bring about novelty, which by definition cannot be predicted (Ellis, 2005), as it is the result of dynamic self-organizational properties that cannot be established a priori (Davids et al., 2003; Lames, 2003). Furthermore, all predictions have to deal with the sensitivity of systems to the initial conditions (Aicinena, 2013), whereby the slightest differences may result in extremely amplified divergences after a period of time (Cubitt et al., 2015). In this respect, even so-called non-linear periodization is linear in its prediction of outcomes (Denison and Scott-Thomas, 2011), as it uses a sum-of-components approach, which is incompatible with the nature of complex systems (Ellis, 2005).

In light of these concerns, theoretical criticisms have been addressed to research on periodization of training. For example, it has been suggested that research appears to bypass a core principle of training theory and methodology: the divergence between external and internal load (Bailey and Pickard, 2010; Scanlan et al., 2014). Indeed, considerable inter-individual differences in response to training, nutrition, supplementation, and medicine have been well established in sport (Kenney et al., 2012). It is known that some persons are non-responsive to certain types of stimuli, e.g. pulmonary rehabilitation (Stoilkova-Hartmann et al., 2015), cardiac resynchronization therapy (Auricchio and Prinzen, 2011), viral infections (Perng and Chokephaibulkit, 2013), and use of antidepressants (Kudlow et al., 2014), among others. Even within those who are responders, there is a wide range of variation, from low- to high-responders (e.g., Perng and Chokephaibulkit, 2013; Stoilkova-Hartmann et al., 2015).

The existence of non-responsiveness is extensive to training regimes. Following a protocol of live-high, train-low training at 1650 m, Hamlin et al. (2011) found that some athletes markedly changed the sympathetic-to-parasympathetic ratio, while others were non-responsive, i.e., their autonomic nervous systems’ activity did not change after the training protocol. Non-responders to altitude training, specifically live high-train low protocols, comprise nearly 50% of the tested population (Paula and Niebauer, 2012). In a study with youth football players, Faude et al. (2013) found that 40% of players subjected to High-Intensity Interval Training did not change their individual anaerobic thresholds. With regard to resistance training, extensive variation in responses have been verified. Subjects range from low-responders to high-responders when changes in muscle size and strength are considered, suggesting we need to further focus research further on inter-individual variation in responses (Ahtiainen et al., 2016; Fisher et al., 2014; Garcia et al., 2016).

Although proponents of periodized programs have underlined the need to respect inter-individual variations...
in accommodation to a given training stimulus (e.g., Bompa, 1999; Mujika, 2007), research on periodization appears to have ignored such variations (Kiely, 2012; Lames, 2003). Moreover, intra-individual variation in time has escaped analysis in periodization research (Alcainena, 2013; Kiely, 2012; Lames, 2003). Finally, periodization is being equated with variation, although periodized programs are more than random variations; conversely non-periodized programs can be varied (Harries et al., 2015; Kiely, 2012; Turner, 2011). This leads to errors of analysis due to conceptual equating of what are two distinct concepts.

Systematic reviews and meta-analyses have synthesized empirical findings on the application of periodized training programs. Dantas et al. (2010) conducted a systematic review of 103 papers on periodization, having concluded that the models of Classical Periodization, Accumulation-Transformation-Realization, and Structural Bells were superior to models of Prioritized or Block Practices. However, this systematic review included book chapters, technical papers, and non-accredited web sources. In addition, non-periodized programs were equated with constant volume programs, despite the fact that non-periodized programs can be varied. A meta-analysis was published by Dantas et al. (2011), and the results suggested the superiority of the models of Matveev, Verkhoshansky, and Bompa, in comparison to those of Accumulation-Transformation-Realization and Forteza. Unfortunately, this paper suffers from the same problems as the above mentioned systematic review.

In the meta-analysis conducted by Rhea and Alderman (2004), periodized models presented superior results when compared to non-periodized models with respect to strength and power outcomes, but once again non-periodized programs were equated with non-varied programs. A systematic review with meta-analysis (Harries et al., 2015) has shown no differences in the effectiveness of linear versus undulating periodization on strength. The authors stated that their results suggested that variety and novelty in training were the important factors, whereas the specific type of variation might have been more relevant. Hence, variation alone and not periodized variation might be the key.

Finally, when applying training protocols, it is important to control for nutrition, supplementation and medicine, as these factors may influence outcomes (Kenney et al., 2012). Overall, nutrition strategies and timings, as well as supplementation profoundly impact the outcomes of training programs (e.g., Helms et al., 2014; Perez-Schindler et al., 2014; Pyne et al., 2014). Protein supplementation, for example, enhances hypertrophic gains, being more effective than resistance training alone (Pasiakos et al., 2015; Phillips, 2016). Medicine also interferes with responses to exercise, such as the combined use of statins and exercise training (Deichmann et al., 2015). Even caffeine intake can alter the responses to training stimuli (Kenney et al., 2012). As such, well-designed studies should attempt to at least report on some of these parameters.

We propose, therefore, to conduct a comprehensive review of how empirical research on training periodization with human subjects has been performed. In particular, we aim to answer the following research questions: (1) Are the concepts of periodization and variation actually being used as synonyms? (2) Is research on periodized exercise programs actually testing the direction, timing, and magnitude of adaptations? (3) What time-frames are being considered in such research (e.g., short-, medium-, and/or long-term)? (4) What dimensions of load are being investigated? And (5) Are confounding factors being declared?

Methods

Search criteria

The search was conducted in late February 2016 on the following databases: EBSCO + SportDISCUS (specifically selecting Academic Search Complete, CINAHL Plus with Full Text, MedicLatina, MEDLINE with Full Text, PsycINFO, and SPORTDiscus with Full Text), PubMed, Scielo, Scopus and Web of Science. No limitations were imposed concerning date of publication, and in press papers were included. Search and retrieval of papers was conducted by two of the researchers independently and simultaneously. The study was conducted in accordance with the principles of the Helsinki Declaration and was formally approved by the ethics committee of the Faculty of Sports of the University of Porto.

Inclusion criteria

Only original empirical articles published in peer-reviewed journals with a full manuscript available were included. Boolean operators were used for the searches. The titles had to include the terms “periodization” OR “periodized”; orthographic variations such as “periodization” and “periodized” were considered and accepted. The operator OR conjugated these words in the title with “exercise” OR “sport” OR “training” in the title or abstract. As these search engines automatically translate titles written in other languages, articles were included if a full manuscript had been written in one of the following languages: English, French, Italian, Portuguese or Spanish. Only studies with human participants were included. Duplicate papers (i.e., emerging in several databases or more than once in the same database) were counted only once. Overall, 118 papers were retrieved in this stage. Again, two of the authors independently conducted the whole process. Where disagreements were found, a reanalysis of the search and inclusion criteria was implemented.

Exclusion criteria

Twenty-six papers were excluded because they did not actually analyze periodization (e.g., plyometric program vs. weight training program). Observational studies ($n = 31$) were excluded, as with such designs it is always possible to question: i) if these results have been derived from sport-specific practice and not from the experimental protocol? and ii) if the experimental protocol actually have had a detrimental effect upon what would be the benefits of the sports-specific practice? Case studies ($n = 3$) were also excluded, as they were observational reports. 
with very small samples (one to three participants). Within experimental papers, trials that were both non-randomized and non-controlled were also excluded ($n = 4$) as there was no accurate way to interpret data. Four controlled but non-randomized studies were excluded as biases in sample distribution could influence the outcome. One randomized controlled trial in a clinical setting was excluded due to its extremely short duration (one week). Three randomized controlled trials (RCTs) were excluded due to lack of proper information about how the control group was performing. Four RCTs were excluded since they only presented one experimental group versus one non-training group. Two authors completed the exclusion stage independently. A third author operated to analyze cases where disagreement might exist. Figure 1 synthesizes the information pertaining the whole process.

**Sample**
A total of 42 original empirical papers were selected: 24 randomized trials (RTs) (Bartolomei et al., 2014, 2015; Buford et al., 2007; Foschini et al., 2010; Franchini et al., 2015; González-Ravé et al., 2007; Herrick and Stone, 1996; Kok et al., 2009; McNamara and Stearne, 2010; Miranda et al., 2011; Monteiro et al., 2009; Pacobahyba et al., 2012; Painter et al., 2012; Prestes et al., 2009a; 2009b; Ramalingam and Yee, 2013; Rhea et al., 2002; 2003; Ronnestad et al., 2014a; 2014b; 2016; Sauer et al., 2014; Schiotz et al., 1998; Zourdos et al., 2016) and 18 randomized controlled trials (RCTs) (Abt et al., 2016; Ahmadizad et al., 2014; Apel et al., 2011; Baker et al., 1994; DeBeliso et al., 2005; Esteve-Lanao et al., 2008; Kell, 2011; Kraemer et al. 2003; Lacordia et al., 2011; Lima et al., 2012; Moir et al., 2007; Moraes et al., 2013; Perez, 2013; Simão et al., 2012; Souza et al., 2014; Spineti et al., 2013; Willoughby, 1991; 1993). Randomized controlled trials provide, arguably, the highest-quality evidence to support evidence-based practice (Cumming, 2014), but what some researchers considered to be control groups, others reported as being experimental groups, hence the maintenance of RTs in our review.

**Variables**
First and foremost, how non-periodized programs were conceptualized was analyzed. Specifically, it was our intent to understand whether authors equated non-periodized approaches with constant approaches, thereby mistaking periodization with variation. Secondly, predictions were analyzed regarding direction, timing, and magnitude of expected adaptations.

With regard to study design, dimension of load was verified (e.g., strength), as well as the length of the protocol (short-term: 4 to 12 weeks; medium-term: 13 to 18 weeks; long-term: >18 weeks). We also aimed at analyzing the steps taken to control for nutrition, supplementation and/or medication. In addition, data analysis was scrutinized with respect to dispersion analysis in the discussion of the findings. Finally, reporting of effect sizes was also considered.

**Summary of PICOS process**
Population consisted of human subjects. Intervention was a narrow application of periodized training protocols. Comparison was between different periodized approaches or between periodized and non-periodized approaches in any training protocol. Outcomes were free to vary; hence they did not constitute an exclusion criterion. Study design was limited to randomized trials or randomized controlled trials.
Reliability of analysis
To ensure the reliability of the analysis, two of the authors went through the entire process of search and retrieval of papers fully, as well as the coding of the data. At the end of the process, the two authors carefully compared their coding tables; any inconsistency in coding was re-checked by rereading the paper to avoid unwanted mistakes. Across the whole process, a third author verified the different stages in order to account for and resolve any inconsistencies and disagreements.

Results

Conceptual issues
Data showed that the concepts of periodization and variation were being used as synonyms, as all papers equated periodized programs with pre-arranged varied programs, while non-periodized programs were equated with constant (i.e., non-varied) programs.

Furthermore, no paper ventured into predictions of direction, timing or magnitude of the adaptations. Therefore, the effectiveness of load management was not actually tested in any of the papers.

Issues with study design
Our data has shown that researchers are not considering global perspectives of performance. The physical dimension represented 95.2% of the research process \((n = 40)\). Within it, strength represented the main focus, corresponding to 57.1% \((n = 24)\) of research, followed by strength and endurance \((14.3\%, n = 6)\), endurance \((9.5\%, n = 4)\), strength and power \((4.8\%, n = 2)\), strength, power and flexibility \((2.4\%, n = 1)\), strength, power, endurance and agility \((2.4\%, n = 1)\), power \((2.4\%, n = 1)\), and power and speed \((2.4\%, n = 1)\). Only one empirical paper \((2.4\%)\) focused on technique, while another focused on physical and tactical aspects. Research on periodization of psychological factors was absent.

With regard to the length of the experimental protocols, only two papers \((4.8\%)\) presented research lasting 24 weeks or above, meaning that the vast majority of the research lasted less than six-months. No paper extended further than nine months in duration. Most studies were situated between four and 12 weeks of duration \((73.8\%, n = 31)\), with the 12-week mark being predominant \((42.9\%, n = 18)\). Research lasting between 13 and 18 weeks comprised 19.0% of the sample \((n = 8)\).

Respecting potentially confounding factors (i.e., nutrition, supplementation, and medication), 54.8% of the studies did not report any information on these topics. One study \((2.4\%)\) provided dietary guidelines to the participants, but did not control compliance to such guidelines. One study used dietary logs \((2.4\%)\). Four investigations \((9.5\%)\) presented guidelines for the hours preceding testing sessions, but not during the experimental program as a whole. Four papers \((9.5\%)\) prohibited the utilization of nutritional supplementation or ergogenic aids. Four papers did the same while also providing dietary guidelines, whereas two papers \((4.8\%)\) combined prohibition of nutritional supplementation or ergogenic aids while controlling for diet or using nutritional therapy. One study \((2.4\%)\) provided supplementation to all groups as part of the protocol, and another delivered protein supplementation with dietary logs being kept. One study reported that the participants were not taking medication and were non-smokers. Overall, only seven papers attempted to control or, at least, report two of the three factors (e.g., nutrition plus supplementation). No paper controlled all three factors.

Issues with data analysis
Data dispersion (i.e., data on variation) was presented, usually in the form of standard deviations, but the fact remains that discussions were focused around central values and mainly in between-groups differences, while not exploring within-group variations in responses. One study did mention that the experimental group presented a large standard deviation, denoting large variations in the responses to training, even if the sample was quite homogeneous (Moir et al., 2007). Another study underlined the fact that standard deviation was higher after the intervention (Simão et al., 2012). In both cases, these issues were explored no further. No paper devoted any attention to possible low-responders, high-responders, or non-responders.

Additionally, only nineteen papers \((45.2\%)\) have reported properly the magnitude of observed effects. More than half of the sample \((n = 22; 52.4\%)\) did not calculate effect sizes, while one paper \((2.4\%)\) provided a very incomplete report.

Discussion
Periodization of training and exercise is a widespread practice in both competitive and health settings (Issurin, 2008; Naclerio et al., 2013). However, this has been subject to various theoretical critiques (e.g., Aicinena, 2013; Denison, 2010; Denison and Scott-Thomas, 2011; Kiely, 2012). The aim of this comprehensive review was to analyze empirical research on training periodization with human subjects. A set of conceptual and methodological problems was identified in this work.

Conceptual issues
Our data reveals that the concepts of periodization and variation are being used interchangeably in research, while they actually represent two distinct constructs. Mere variation does not provide the basis for a periodized program (Kiely, 2012), as acknowledged by mainstream definitions of periodization (e.g., Harris et al., 2015). Instead, variation must be controlled in order to pursue specific timelines for peak performances and avoidance of overtraining (Fleck, 2008; Turner, 2011); only when these criteria are met can we properly refer to an exercise program as being periodized. Conversely, non-periodized programs can be varied. To our knowledge, no empirical paper, systematic review or meta-analysis has even addressed this major issue.

Furthermore, and perhaps even more surprisingly, given the self-proclaimed goal of periodization of being able to predict periods of peak performance while avoiding overtraining (Fleck, 2008; Naclerio et al.,
Turner, 2011), is the fact that no empirical paper attempted to make any predictions concerning the direction, magnitude, and especially the timings of the adaptations. In sum, research on periodization has avoided putting forth testable predictions.

**Issues with study design**

With regard to study design, the ‘physical’ load is predominant (40 out of 42 papers), as if the technical, tactical, and psychological dimensions of load were not part of the physical load! The predominance of unidimensional designs, mainly focused on the “physical” aspects of training, configures itself as a by-product of a biased conceptual understanding of load. Paradoxically, in many sports tactical and technical aspects are deemed more relevant than the “physical” factor (Williams and Hodges, 2005).

Another problem with study design concerns the length of the experiments. Most sports have evolved in the direction of long seasons, demanding that performance is kept at high levels over several months in a row (Gamble, 2006; Mujika, 2007). In contrast, most empirical studies conducted on periodization were short-term (73.8%) or medium-term (19.0%). However, it is known that results from short-term research on periodization should not be transposed into longer-term periods (Fleck, 2008; Harries et al., 2015). In our systematic review, no paper surpassed 9 months in duration.

Additionally, relevant confounding factors such as medication, supplementation, and nutrition were only marginally accounted for in the analyzed papers: more than half of the papers did not report any information on these issues, and no paper reported information on all three factors. This is hardly credible, as all these factors are prone to impact greatly on the effects of any training program (e.g., Deichmann et al., 2015; Kenney et al., 2012; Phillips, 2016; Pyne et al., 2014).

**Issues with data analysis**

In relation to data analysis, variation in response to training programs is being ignored because of averaged values and considerations. Our results highlighted that only two papers made some reference to variations in responsiveness (Moir et al., 2007; Simão et al., 2012). Effectively, empirical research on periodization has neglected a core principle of training theory and methodology: inter- and intra-individual variation in response to training programs (Bailey and Pickard, 2010; Kenney et al., 2012; Scanlan et al., 2014). It has also ignored a wide body of research showing that variations in responsiveness to training programs are the norm (e.g., Ahtiainen et al., 2016; Faude et al., 2013; García et al., 2016; Paula and Niebauer, 2012). This presents a major challenge to the premises behind periodization, but research on periodization has remained silent with regard to such topics.

Effect sizes were also often unreported or incompletely reported, an awkward option since statistically significant effects may not translate into practical significance (Nuzzo, 2014; Winter, 2008). Effect sizes help to answer the question if an intervention works, and how well (Winter et al., 2014). Therefore, it would be expected that researchers would want to quantify the impact of their training programs, but that is not always the case.

**Overall analysis**

Training is an ambiguous and unpredictable activity (Barker and Bailey, 2015), but often coaches try to pre-establish an orderly sequence of events and effects (Denison, 2010). Rationalistic conceptualizations of coaches’ practices presuppose a certainty of outcome that is incompatible with the intrinsic indeterminacy of practice (Jones and Corshy, 2015; Jones and Wallace, 2005). Broad-scope metatheories, or worldviews, inform scientific research programs (Overton, 2015), and are frequently so deeply ingrained that researchers may not even perceive that they are abiding by its postulates (Overton, 2014; Thelen and Bates, 2003). In the case of sports periodization, the problem is more likely to derive from the larger prediction-like metatheory that has been criticized in fields such as economics (Hendry and Mizon, 2014), history (Arendt, 1998), and talent identification (Abbott et al., 2005; Harder et al., 2014). Narratives suitable for training processes should emphasize change, contingency, context, improvisation (Denison, 2016).

**Conclusion**

This comprehensive review of empirical research on training periodization has shown that: i) the concepts of periodization and variation are being used interchangeably; ii) predictions concerning the direction, timing and magnitude of adaptations are not actually being tested; iii) analyses have been mostly unidimensional, focusing almost exclusively on the ‘physical’ aspects of performance; iv) long-term empirical papers are non-existent; v) confounding factors such as medication, nutrition, and supplementation are not being properly reported and controlled; and vi) data interpretation is being compromised by persistently ignoring inter-individual variation in responsiveness to experimental protocols.

**References**


**Key points**

- Periodization is considered a core concept of training.
- However, conceptual and methodological critiques have risen.
- We therefore comprehensively reviewed randomized and randomized trials applying periodized protocols to human subjects.
- Overall, the concepts of periodization and variation are being used interchangeably, which represents an intellectual mistake with implications for how we interpret the results of the studies.
- Additional methodological shortcomings make current research on periodization largely unreliable.

---

**AUTHOR BIOGRAPHY**

**José AFONSO**

Employment  
Professor, Faculty of Sport – University of Porto, Portugal  
Degree  
PhD  
Research interests  
Coaching and performance analysis  
E-mail: jneves@fade.up.pt

**Pantelis T. NIKOLAIDIS**

Employment  
Professor, Department of Physical and Cultural Education – Hellenic Army Academy, Athens, Greece  
Degree  
PhD  
Research interests  
Performance Analysis, especially through Sports Physiology.  
E-mail: pademil@hotmail.com

**Patrícia SOUSA**

Employment  
Researcher, Faculty of Sport – University of Porto, Portugal  
Degree  
MSc  
Research interests  
Periodized approaches to training.  
E-mail: patricia.spsousa@hotmail.com

**Isabel MESQUITA**

Employment  
Professor, Faculty of Sport – University of Porto, Portugal  
Degree  
PhD  
Research interests  
Performance analysis and coaching.  
E-mail: imesquita@fade.up.pt

José Afonso  
Centre for Research, Formation, Innovation, and Intervention in Sport. Faculty of Sport – University of Porto, Portugal