• Sport nutrition is one of the fastest growing and evolving disciplines of sport and exercise science with research ranging from the discovery of novel nutrient sensitive cell signaling pathways to the effects of sports drinks on ratings of perceived effort.
• Sport nutritionists often utilize both classic and contemporary literature to help guide their applied practice, ultimately hoping that the research can help to fuel winning performances.
• Step one in the translation of research to practice should always be a well-structured critique of the translational potential of the existing scientific evidence.
• We present an operational framework (the paper to podium matrix) providing a checklist of criteria to prompt the critical evaluation of performance nutrition related research papers.
• It is the combination of boldness of reform (i.e., innovations in research) and quality of execution (i.e., ease of administration of practical solutions) that is most likely to deliver the transition from paper to podium.

INTRODUCTION
The idea of testing nutritional interventions dates back many centuries. However, sports nutrition as both an academic and practical discipline, has stemmed mainly from research performed in the late 1960s, following a series of seminal studies examining the effects of muscle glycogen on exercise capacity and performance (Bergstrom & Hultman, 1966 a,b; Bergstrom et al., 1967; Hermansen et al., 1967). Since these landmark articles, the field of sports nutrition has developed significantly with studies examining the effects of nutrient availability and ergogenic aids for modulation of performance, recovery, training adaptation and body composition (Thomas et al., 2016). Although the practitioner’s goals should always be the delivery of research-informed practice, there is often a required balance struck between waiting for the peer-reviewed randomized controlled trials (RCTs) and subsequent meta-analysis, and the early adoption of novel methods in an attempt to gain a competitive advantage (Coutts, 2016). The sport nutrition discipline may be viewed as highly confusing and contradictory, with numerous real-world examples where research findings have been misconstrued to inform applied practice. It is imperative for the growth and development of the sports nutrition area that a well-structured critique of available peer-reviewed scientific evidence is conducted prior to use in practice.

The aim of this sports science exchange article is to provide the sports nutritionist with a time-efficient operational framework to critically evaluate the translational potential of applying research to practice (Close et al., 2019). The aim of this framework is not to critique the general methodology of research, but instead appraise the direct and applied applicability of research through the use of our Paper to Podium (P2P) Matrix (Figure 1.)

RESEARCH CONTEXT
As the area of sports science continues to grow, research has moved toward investigation of the underpinning mechanisms behind adaptation to exercise and nutritional interventions using molecular biology techniques (Close et al., 2016). Cell culture and rodent models are often utilized to investigate physiological responses (Gomez-Cabrera et al., 2008; Holland et al., 2016; Owens et al., 2015); however, findings may or may not have application to the exercising human during whole body exercise. While identifying suppression or activation of a specific signaling pathway is likely to yield new insights and research questions, this does not mean that an athlete’s training or nutritional program should immediately change accordingly. It is important to not only consider the in vitro model/technique, but also whether the study was conducted using human cell types as opposed to rodent, as these nonhuman cellular investigations are often designed to examine the potential mechanisms of actions of adaptation and were never intended to be directly translated into practice. It is crucial to assess if the trial has ecological validity (i.e., the methods somewhat replicate the conditions expected in the real-world setting) when assessing the translational ability of the paper. Initial and careful consideration of the research context therefore provides the platform to further evaluate the translational potential to a specific sporting situation.

PARTICIPANT CHARACTERISTICS
Currently, there is often a large amount of ambiguity when describing physiological characteristics and the general training status of participants. Often descriptions such as ‘well-trained’, ‘elite’ and ‘world-class’ are used interchangeably, cited incorrectly, and sometimes provided without any detail regarding habitual training or physiological health and fitness measures. Jeukendrup et al. (2000) initially delivered
an objective standards framework for categorization of cyclists depending upon qualitative (e.g., ranking, race days per annum, training frequency) and quantitative (\(\text{Watt}_{\text{max}}\), maximal oxygen uptake (\(\text{VO}_{2\text{max}}\)) economy) criteria. The idea of defining habitual and physiological standards such as this may be pertinent in all endurance sports; however, within team sport scenarios, it may be practical to simply classify subjects as amateur, semi-professional or professional depending on the grading of competition they compete in. Authors should describe participants using as much quantifiable data as possible and where this data is not cited, the results must be treated with some degree of caution given the vast differences in response to exercise between untrained and trained participants (Bergman & Brooks, 1999). In the context of performance, it is therefore possible that the performance-enhancing effects of a specific intervention (e.g., beetroot juice) is negated in trained vs. less trained participants (Willerson et al., 2012). As such, the efficacy of any nutritional intervention should be investigated in the specific population of which the intervention is intended to be used in practice. In this regard, good examples include studying the efficacy of ketogenic diets in Olympic-standard race walkers (Burke et al., 2017) or work from our group that examined rugby-game glycogen utilization taken from professional players during a competitive fixture (Bradley et al., 2017).

**DIETARY & EXERCISE CONTROLS**

Despite published guidelines for standardization within research (Jeacocke & Burke, 2010), literature often fails to objectively control habitual diet or report any dietary controls employed. Indeed, there are also discrepancies present surrounding the actual delivery of dietary controls, ranging from subject self-reporting, researchers freshly preparing, and/or meal preparation companies preparing and delivering. In addition, it may be useful to also report and standardize exercise during the days leading to the experimental trial, especially within situations that may lead to differences in pre-exercise muscle glycogen availability and thus altered performance. There are, of course, advantages and disadvantages to many of the common dietary standardization methods outlined above, including cost and ease of intervention but also the ecological validity to the research participants. Fasting is often utilized as a control for dietary intervention studies given that feeding around exercise can significantly alter metabolic responses pre-, during, and post-exercise (Bartlett et al., 2013; Horowitz et al., 1997; Lane et al., 2013; Widrick et al., 1993). However, it is unlikely that elite athletes would compete or, indeed, undertake consecutive aerobic and resistance training sessions in the fasted state or without energy intake between sessions; again, caution should be taken when translating research from studies on fasted participants to athletes intending to compete in the fed state. Finally, practitioners should be aware that some studies purposefully restrict the nutrient of interest, e.g., nitrate (Lane et al., 2014) or polyphenols (Bell et al., 2014) in an attempt to ensure a consistent pretrial concentration of the substance of interest. While this approach is scientifically sound, some
caution should be exerted when applying the results given that the athlete in the real world may actually consume a diet naturally rich in the compound of interest. Taken together, we therefore recommend that practitioners carefully evaluate research designs and dietary protocols in relation to the nutritional practices, training loads and training organizational practices that are inherent to their specific sport. And potentially, future study designs should consider utilizing conditions considered to be “best nutritional practice” within their real-world context to assess the “true” magnitude of intervention.

VALIDITY & RELIABILITY OF EXERCISE PROTOCOLS & PERFORMANCE TESTS

The ecological validity, reliability and real-world context of exercise models are often overlooked within research, which alters the ability to interpret and translate the data directly into practice. One such example of this is the one-legged knee extensor model (Andersen et al., 1985), which has been used extensively in exercise metabolism research to evaluate local control of muscle metabolism and adaptation to exercise training. From a mechanistic perspective, this model is advantageous, as examination of responses can be conducted in the same individual with the resting leg as a control. However, the single leg modality fails to accurately replicate exercise in vivo, given that these models have a limited muscle mass, low cardiac function and a reduced hormonal response (Helge et al., 2007). It is noteworthy that dose-response studies evaluating the optimal protein dose to stimulate muscle protein synthesis suggest that the absolute protein dose is effectively doubled (i.e., 20 to 40 g post-exercise protein feeding) when using whole body resistance training protocols (Macnaughton et al., 2016) vs. unilateral exercise protocols (Moore et al., 2009; Witard et al., 2014).

In addition to ecological validity, there is also the requirement to carefully consider the reliability of any exercise performance tests as well as the inclusion of any familiarization trials. Within the context of endurance type performance tests, it can also be debated as to whether the participant should have access to any internal (e.g., heart rate) or external (e.g., power output, running velocity) cues during testing (Edwards & McCormick, 2017) as well as the validity of time trial vs. exercise capacity tests (Currell & Jeukendrup, 2008; Karsten et al., 2018). For example, in the case of professional road cycling, it could be argued that the “true” effects of any nutritional intervention or ergogenic aid should always be evaluated with access to external cues, given that riders have continual access to power meters, heart rate data and verbal feedback from accompanying support staff. Additionally, the use of time-trial and exercise capacity tests could both be considered as valid performance measures, given that both situations do occur in the form of designated time-trial stages and the ability to respond to “attacks” on mountain climbs, respectively. We need to remember, the controlled, calm and temperature-controlled laboratory environment is usually never representative of the elite sporting arena. Would the effects of caffeine on motivation be the same in a calm laboratory environment compared with walking out at Twickenham for a Rugby World Cup final in front of 80,000 spectators?

DATA ANALYSES & PRESENTATION

Often, one of the most contradictory components of performance nutrition related research is the way in which data has been analyzed and presented. In the realm of applied performance research, there has been a recent trend to adopt the approach of magnitude-based inferences (Batterham & Hopkins, 2006) as opposed to more traditional probability-based testing and, as such, it is important that applied practitioners are familiar with the advantages and disadvantages of both methods. Importantly, researchers should provide a clear rationale and justification of the chosen sample size (often accompanied by power calculations) and reasons for their choice of statistical analysis. Finally, the approach to data presentation can also greatly influence how the results are evaluated and interpreted by readers. For example, presentation of group means and standard error (as opposed to standard deviation) does not provide a true representation of the variability between subject responses, but researchers often choose to represent variability using standard error (especially in graphical format) for cosmetic reasons (Morton, 2009).

Given that practitioners usually pursue the application of interventions with individual athletes, richer evaluations of data should be made, and individual differences plotted and visualized where possible, especially when the research was conducted on a small sample. For example, in a recent study from our laboratory examining the effect of muscle glycogen availability on endurance capacity, we observed that mean exercise capacity was increased by 60 min with high vs. moderate pre-exercise glycogen concentration (i.e., 600 vs. 300 mmol.kg⁻¹ dry weight) (Impey et al., 2016). Nonetheless, the individual magnitude of increase in time to exhaustion ranged from 4 min to 113 min. Clearly, evaluation of individual responses can be lost in translation where only group means are presented.

PRACTICAL APPLICATIONS

• Consider the translational potential and context in which the data are collected.
• Investigate the participant characteristics from studies to decide whether the research has direct applicability to your own context.
• Study the research design fully and consider whether it actually was the nutritional intervention that caused an effect.
• Decide whether the research used adequate dietary and/or exercise controls to confidently attribute differences to their intervention of choice.
• Search for any validity and reliability issues within the chosen exercise protocols and performance tests, paying particular attention to the environment in which the study was conducted.
• Reflect on how the data analyses were conducted, how the data were presented, and look for individual differences and standard deviation values.
**Figure 1.** The Paper to Podium (P2P) Matrix: an operational framework to evaluate the translational potential of performance nutrition research. For each category, assess the manuscript on the -2 to +2 scale and then add together the total score. A positive score suggests that the paper has significant translational potential while a negative score suggests some degree of caution should be taken when trying to directly translate the findings to the applied world. It must be stressed that this matrix is by no means an assessment of the quality of the research per se; rather, it is purely an assessment of the translational ability of the manuscript.

<table>
<thead>
<tr>
<th>Category</th>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th>+1</th>
<th>+2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Participants</strong></td>
<td>Levels of participants not reported.</td>
<td>Inappropriate training status or age for the context required.</td>
<td>Inappropriate training status (with defined criteria) although in required age group.</td>
<td>Close to relevant training status and age (with defined criteria).</td>
<td>Relevant training status and age (with defined criteria).</td>
</tr>
<tr>
<td><strong>Research Design</strong></td>
<td>No control group. No blinding of intervention. No consideration of sample size.</td>
<td>Control group but no blinding. No consideration of sample size.</td>
<td>Randomised control trial (RCT) with repeated measures or matched groups design. Control group but no blinding. No sample size calculations but similar to previous research.</td>
<td>RCT with repeated measures or matched groups design. Single blind placebo controlled. Sample size calculated.</td>
<td>RCT with repeated measures or matched groups. Double blind placebo controlled. Sample size calculated.</td>
</tr>
<tr>
<td><strong>Control</strong></td>
<td>No reference to dietary or exercise control cited (but self-reported) with no supported data.</td>
<td>Methods of dietary and exercise control cited (but self-reported) with no supported data.</td>
<td>Methods of dietary and exercise control cited (but self-reported) with supported data.</td>
<td>Dietary provision provided with no supporting data. Exercise control cited. No replication to real-world context.</td>
<td>Dietary provision provided with supporting data. Exercise control cited. Representative of real-world context.</td>
</tr>
<tr>
<td><strong>Validity &amp; Reliability</strong></td>
<td>No familiarisation trial or reliability data and measurement tool error. Exercise protocol not representative to real-world context.</td>
<td>Familiarisation trial. No reliability data or measurement tool error. Exercise protocol not representative to real-world context.</td>
<td>Familiarisation trial. Reliability data and measurement tool error. Exercise protocol not representative to real-world context.</td>
<td>Familiarisation trial. Reliability data and measurement tool error. Exercise protocol representative but laboratory based.</td>
<td>Familiarisation trial. Reliability data and measurement tool error. Exercise protocol representative to real-world.</td>
</tr>
<tr>
<td><strong>Data Analytics</strong></td>
<td>Analytics not reported or performed.</td>
<td>Analytics reported but limited to descriptive statistics.</td>
<td>Analytics reported. Appropriate significance or magnitude-based inference (MBI) tests.</td>
<td>Analytics reported. Appropriate significance or MBI tests. Effect sizes included.</td>
<td>Analytics reported. Appropriate significance or MBI tests. Effect sizes included. Presentation of individual responses to treatment intervention.</td>
</tr>
<tr>
<td><strong>Timing</strong></td>
<td>Not age appropriate. Time for dosage not optimal. Time from major complication insufficient.</td>
<td>Age appropriate. Time for dosage not optimal but could be effective. Time from major complication insufficient.</td>
<td>Age appropriate. Time for dosage is not optimal but could be effective. Time from major competition insufficient.</td>
<td>Age appropriate. Time for dosage is not optimal but could be effective. Time from major competition sufficient.</td>
<td>Age appropriate. Time for dosage is considered optimal. Time from major competition sufficient.</td>
</tr>
</tbody>
</table>
SUMMARY
Although relatively simple in concept, the translation of research to practice is not always a straightforward process. Indeed, elite sport is dynamic, unpredictable and often chaotic, none of which can be interpreted by a two-way ANOVA or predicted from the controlled laboratory environment. Despite the continual pursuit and often-impatient demands for the latest-winning edge, we consider that the starting point for the research-informed practitioner should always be the critical evaluation of the translational potential of the available scientific evidence. Put simply, we must look beyond the abstract, the 140-character tweet and latest infographic in order to truly evaluate the scientific rigor and translational potential of performance nutrition related research studies. Utilization of the P2P Matrix (Figure 1) may help practitioners to personally evaluate a research paper, increasing their own confidence in the intervention they are about to implement, which may ultimately result in a more enthusiastic consultation with the athlete and increase the chance of an effective intervention. It is readily acknowledged that the content and indices of such a framework are not exhaustive. Rather, it was the deliberate aim to provide a time-efficient evaluation tool that can be readily applied by practitioners who all too often operate under the intense time constraints inherent to elite sport. Subsequent to the evaluation of existing research, we also encourage practitioners to conduct field-based research (e.g., case reports or small sample size studies) with the same degree of scientific rigor and precision of measurement that is requisite of RCTs. Ultimately, it is the combination of boldness of reform (i.e., innovations in research) and quality of execution (i.e., ease of administration of practical solutions) that is most likely to deliver the transition from paper to podium.

REFERENCES


