

Editorial**EDITORIAL: The Use of Big Data to Improve Human Health - How Experience from Other Industries Will Shape the Future**Timothy E Hewett, PhD^{1, a}, Kate E Webster, PhD²¹ Hewett Global Consulting, Rochester, MN, ² La Trobe University, Melbourne Australia

Keywords: research, data, biostatistics

<https://doi.org/10.26603/001c.29856>

International Journal of Sports Physical TherapyVol. 16, Issue 6, 2021

The article in the current issue of the International Journal of Sports Physical Therapy entitled “Using Big Data to Improve Human Health: How Experience from Other Industries Will Shape the Future” is both highly relevant and important to the readership of IJSPT as it will challenge our thinking of traditional clinical and research paradigms and approaches and present new possibilities.

The goal of traditional medical and biostatistical inquiry has been to determine cause and effect, and hence predictive relationships between independent and dependent variables that are hypothesized to be important to human health. Early twentieth-century biostatisticians (see the great Ronald Fisher, for example) therefore established randomized control trials (RCTs) and data science approaches that aimed to compare the differences between test groups (intervention versus “control”). These traditional medical biostatistical approaches compare the differences between intervention versus “control” groups as the numerator and the measurement error in the denominator of the equations. Hence, the focus of these comparisons relies on highly precise measurements of both the differences between groups and the errors between groups, which are often quite challenging to determine. More recent data science approaches that utilize Artificial Intelligence schemes such as Machine Learning and Neural Networks do not require as much precision as they model off big data sets and rely on Big Data inputs to develop predictive algorithms.

It was purported by these early biostatisticians that cause and effect could only be determined with medical biostatistical approaches that utilized randomized controlled trials (RCTs). However, subsequently clinicians together with biostatisticians have attempted to determine cause and effect relationships using clinical cohort studies. Association and prediction are the key concepts that are sought

for and utilized in these analyses. Obviously, these cohort studies can help us to determine the association, but can these studies actually be utilized to determine causation and to develop predictive algorithms? For example, clinical cohorts are regularly used to draw cause and effect relationships between musculoskeletal parameters such as body mass, and musculoskeletal injury, such as knee, ankle or shoulder, risk. This is especially the case for clinical questions in which RCTs would be unethical or impossible to conduct. Validation of these findings from clinical cohorts then become of crucial importance.

Studies of the reliability and validity of these cohort findings become necessary. These studies can be highly challenging, especially with limited datasets. There remains a substantial gap between these approaches and the determination of cause and effect and predictive relationships between these limited numbers of independent and dependent variables. Many important questions are raised in the article in this issue of IJSPT, which include: How does the biomedical statistical approach compare to these newer “Big Data” analyses? How can Artificial Intelligence approaches help us to accomplish these predictive goals with greater reliability and validity? Can we combine traditional and Machine Learning approaches for the determination of both cause-and-effect relationships and enhanced predictive capacity that can lead to markedly improved clinical outcomes? As the authors of this paper point out, the convergence of the two disparate approaches will be key to future advancements in biomedical research and its clinical utility.

Submitted: October 15, 2021 CST, Accepted: November 07, 2021 CST

^a Hewett is Chairman and Webster is a member of The Scientific Advisory Board of Sparta Science, a data science company.

Corresponding author:

Timothy E. Hewett

Tim.Hewett@gmail.com

Hewett Global Consulting

1516 4th Avenue NW, Rochester, MN



This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-NC-SA-4.0). View this license's legal deed at <https://creativecommons.org/licenses/by-nc-sa/4.0> and legal code at <https://creativecommons.org/licenses/by-nc-sa/4.0/legalcode> for more information.