TITLE: Utilizing Training Load and Intensity to Predict Team Performance in NCAA Division I Men's Basketball

AUTHORS: Conor A. Kerr, Sam R. Moore, S. Doug Halverson, Abbie E. Smith-Ryan

BACKGROUND: The use of technology in sports has rapidly increased in recent years, with teams collecting vast amounts of data to help inform decision-making and optimize performance. PURPOSE: To utilize practice and game accelerometry data, captured throughout a Division I collegiate men's basketball season, to predict the odds of the team covering the predicted point spread of an upcoming game.

METHODS: The accumulated acceleration loads (AAL), maximum acceleration loads (AL max), mechanical loads (ML), mechanical intensities (MI), and exertions of eight Division I college basketball players of the same team (chosen due to their consistent levels of playing time across the season) were collected via Kinexon's wearable accelerometers (measuring changes in the acceleration of the player's movements) over the course of 161 days, including 33 games. These metrics were matched against games two days prior to, one day prior to, and on the day of the game they were collected. Caesars Sportsbook published predicted margins of victory in terms of point spreads were compared to selected metrics obtained through the accelerometers by a model selection algorithm and multivariate logistic regression model. The algorithm selected the most impactful metrics in predicting whether or not the team exceeded an arbitrary game's point spread using the accelerometer data two days prior, one day prior, and day of the game. The model formulated this prediction using the selection metrics.

RESULTS: Given the entirety of the accelerometer data across all 161 days, the model selection algorithm selected the day-of MI ($\mu = 9.978\,, \sigma = 5.658$), the day prior AAL ($\mu = 414.9, \sigma = 208.1$),

MI ($\mu = 14.44\,, \sigma = 5.419$), and exertions ($\mu = 88.54\,, \sigma = 53.55$), and the two-days prior

ML ($\mu = 1315\,, \sigma = 729.63$), MI ($\mu = 14.41\,, \sigma = 7.106$), and exertions ($\mu = 95.28\,, \sigma = 68.53$) to incorporate into a model predicting the log odds of the team covering the spread in an arbitrary upcoming game. The multivariate logistic regression model created is as follows,

$$\ln\left(\frac{Pr(exceeding\ the\ spread)}{1-Pr(exceeding\ the\ sead)}\right) = 1.102 - 0.042(MI_0) - 0.01(AAL_1) + 0.09(MI_1) + 0.02(Exertions_1) + 0.002(ML_2) - 0.086(MI_2) - 0.015(Exertions_2)$$

The day prior MI ($st.\beta = 1.011\,, p = 0.07388$) and exertions ($st.\beta = 2.208\,, p = 0.0101$) , and the two-days prior ML ($st.\beta = 2.914\,, p = 0.004611$) were positively associated with the log odds of the team exceeding the spread, while the rest of the metrics appear to have negative relationships with these odds. A chi-squared test for a drop in deviance suggested that the probability of a null model having the same explanatory power of the data as the above model is approximately 2.31×10^{-9}.

CONCLUSION: At a more generalized level, the model supports 'short and intense' practices the day before games and 'long and less intense' practices two days before games for covering the predicted spread. The model consistently emphasized the importance of the total volume of a session over its intensity, as seen by the greater standardized effect sizes of the load-related variables (ALL and ML) compared to the intensity-related standardized effect sizes for MI and exertions. This emphasis, along with the characteristics of the overall model suggests that practice planning, particularly in terms of volume, may have a direct link to on-the-court basketball performance.