

Team: MCMC Hammer

Slide 1: Outside Data

To extend our analysis, we thought it would be helpful to include player statistics in order to better understand how the variables we were provided interacted with personal data like age, height, weight, and average points per game. We found a dataset that provided this information, but we did not know how PlayerIDs mapped to actual players. To solve this problem, we used the gps data to identify the substitutes in each world cup game. This matched many of the players with their ID, but some of the players never started out as substitutes, so we identified these players by their tries after approximating the orientation of the field.

Slide 2: Recovery Variable

Our focus for the analysis, was how different variables interact with the amount of time it takes to 'recover' from fatigue after a game. In order to be able to do any kind of inference or analysis we had to be able to quantify recovery. The best way we found to be able to do this was by measuring the number of days it took each player to recover from each tournament. We chose to look at tournaments because all of the games occur either on the same day or in quick succession, and would therefore have no individual recovery times. We used the player's average fatigue as a threshold for recovery. We then looked at individual player's fatigue over time after each tournament and compiled it. We calculated each players average recovery time to be 3.38. Something to note is we discarded player 9, because she overall did not have reliable data, and tended to not have fatigue measurements around tournaments anyway. We also had to drop the Kitakyushu tournament, since it took place so soon after the Commonwealth tournament, and had significantly increased recovery times, as the players were recovering from two tournaments.

Slide 3: Inference

We used multiple linear regression to explore what variables best explained Average Days to Recover from fatigue. We were concerned with recovery from fatigue because the Canadian National Women's Rugby Team wanted to understand and manage player fatigue. By doing so, they could better protect their players from dehydration, injuries, and other harms.

We found several variables that were significant in explaining a player's average recovery time. Although some of the variables cannot be directly changed--Age and Menstruation, for example--the variables that can are worth looking into. Since we found duration to be significant, team managers can take that into account when planning training sessions; however, we found that the session type itself was not significant. It is possible that more detailed data about the session may be produce more insight. Additionally, Player 3 had a high average points scored, and when we removed her, it changed one of the model coefficients from a small, negative number to a larger positive number.