

2º CICLO DE ESTUDOS

MESTRADO EM INFORMÁTICA MÉDICA

Análise de dados biométricos para apoio à gestão de fadiga em atletas de futebol profissional

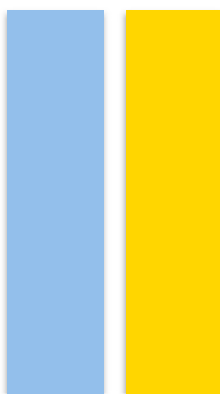
José Luís Ferreira Santos

M

2019

ORIENTADOR: Doutor Miguel Tavares Coimbra

COORIENTADOR: Mestre Carlos Eduardo Maurício Pacheco



Agradecimentos

Começo por manifestar o meu sincero obrigado ao meu orientador, Professor Doutor Miguel Coimbra, por toda a sua disponibilidade, conselhos, correções e anotações, mas acima de tudo, pela sua capacidade de transmitir os seus conhecimentos. Obrigado por todo o apoio prestado na elaboração desta dissertação.

Ao meu coorientador, Professor Carlos Pacheco, pelo desafio lançado e por ter sido a ponte com o clube Futebol Clube de Famalicão. Os seus conhecimentos provaram-se fundamentais na exploração deste tema e por isso e pela sua amizade estou-lhe eternamente grato.

À Neuroexcellence, Unipessoal LDA, pela oportunidade de colaborar nesta área que tanto significado tem para mim.

Ao Futebol Clube de Famalicão, por me receberem e me permitirem colaborar neste ambicioso projeto. Será uma experiência que guardarei para sempre com extremo carinho e de onde trouxe conhecimentos que certamente me irão ser úteis no futuro. As minhas felicitações pelo cumprimento do objetivo principal de subida ao primeiro escalão de futebol nacional.

Um agradecimento especialmente carinhoso, aos meus pais, à Rita, meu pilar e porto de abrigo, à minha irmã e ao meu cunhado, aos meus sobrinhos, Bernardo e Tiago, e aos meus amigos de infância que sempre estiveram e estarão presentes nos momentos mais importantes. Sem a vossa colaboração e compreensão este projeto não teria sido possível.

Por fim, agradeço ao financiamento do projeto NORTE-02-0853-FEDER-037340 - NexTrainer – New Equipment for Advanced Neuromuscular Calf Training and Recovery, cofinanciado pelo programa Operacional Regional do Norte (NORTE2020), através do Fundo Europeu de Desenvolvimento Regional (FEDER).

Sumário

Esta dissertação tem como principal objetivo avaliar a recolha e análise de dados biométricos como a força exercida por grupos musculares chave no contexto do futebol e as movimentações realizadas pelos jogadores em ações de treino e jogo como ferramenta de monitorização de fadiga.

Deste modo, a população de 20 atletas que esteve aos serviços do clube durante a totalidade do período da recolha dos dados foi avaliada sistematicamente em exercícios de adução da anca, extensão do quadríceps femoral e extensão dos isquiotibiais. Durante o mesmo período, foram monitorizados também os movimentos dos atletas em campo, com especial ênfase nas ações de alta intensidade como o número de acelerações e desacelerações, distâncias totais percorridas e distâncias em sprint máximo. Esta observação teve uma duração de 269 dias com início em 13 de agosto de 2018 e conclusão em 9 de maio de 2019.

Após a conclusão do período de aquisição foi processada a base de dados resultante de modo a avaliar a eficácia e validade da monitorização destes parâmetros através de processos de caracterização de mineração de dados. Com o processamento da base de dados foram descartados os dados recolhidos para a flexão do quadríceps femoral e extensão dos isquiotibiais uma vez que a percentagem de valores em falta para ambos os casos rondou os 95%.

Da análise dos valores de Squeeze 45° não foi possível inferir sobre a sua viabilidade como instrumento para monitorização de fadiga dado que os valores observados antes e após o jogo não apresentam diferenças estatisticamente significativas. No entanto, avaliando as movimentações dos atletas durante os jogos e momentos de treino é possível encontrar dados que possivelmente suportam um bom planeamento e gestão do plantel por parte da equipa técnica.

Palavras-chave: Futebol, Fadiga, Data Mining, Análise de Caracterização.

Abstract

This dissertation aims to evaluate the collection and analysis of biometric data such as the force exerted by key muscle groups in the context of soccer and the movements performed by the players in training and game actions as a tool for monitoring fatigue.

Thus, the population of 20 athletes who played at the club during the entire period of data collection was systematically evaluated in hip adduction, quadriceps femoral extension and hamstring extension exercises. During the same period, the athlete's movement in the field was also monitored, with special emphasis on high intensity actions such as the number of accelerations and decelerations, total distances traveled and distances in maximum sprint. This observation had a duration of 269 days starting on August 13, 2018 and ending on May 9, 2019.

After the conclusion of the acquisition period, the resulting database was processed in order to evaluate the effectiveness and validity of the parameters monitored, through data mining characterization processes.

After processing the dataset, the data collected for flexion of the femoral quadriceps and extension of the hamstrings was discarded as the percentage of missing values for both cases was around 95%.

From the analysis of the Squeeze 45° values it was not possible to infer about its viability as an instrument for fatigue monitoring as the values observed before and after the game do not present any significant differences. However, evaluating the movements of the athletes during the games and training moments it is possible to understand that these results can be derived from a good planning and management by the technical team.

Keywords: Football, Fatigue, Data Mining, Characterization Analysis.

Preâmbulo

Durante o meu percurso profissional na empresa Neuroexcellence, empresa que produz e desenvolve sistemas e soluções para contextos desportivos de alto rendimento, fui repetidamente abordado sobre as práticas desenvolvidas e a sua abordagem ótima em desporto de alta competição, mais concretamente na área do futebol profissional.

Fruto destes contactos, surgiu a possibilidade de colaboração com o plantel profissional do Futebol Clube de Famalicão onde a nova equipa técnica do departamento de alto rendimento se propôs a desenvolver uma base de dados cujo principal objetivo seria o apoio à gestão e otimização dos rendimentos do plantel.

Os conhecimentos adquiridos no primeiro ano do mestrado em informática médica aliados à minha experiência profissional possibilitaram-me um papel no desenvolvimento e estudo deste tema.

Após o surgimento desta oportunidade, convidei o Professor Doutor Miguel Coimbra para me orientar neste processo e o Professor Carlos Pacheco, responsável pelo departamento de alto rendimento do FCF para meu coorientador.

Durante toda a época de 2018/2019 tive a possibilidade de adquirir competências em áreas relativas a *data mining* e extração de conhecimento de dados, de modo a contribuir neste projeto.

Considero o tema relevante e como tal, decidi basear o tema da minha dissertação neste projeto.

Com esta dissertação espero contribuir de forma pertinente para a utilização de dados biométricos como ferramenta para a monitorização de fadiga em desporto de alta competição.

Índice

Agradecimentos.....	i
Sumário	ii
Abstract	iii
Preâmbulo	iv
Índice.....	v
Acrónimos	vii
Índice de figuras.....	viii
Índice de tabelas	ix
1. Introdução	10
2. Estado da arte.....	14
2.1 Fadiga muscular	14
2.2 Fadiga periférica.....	18
2.2.1 Fadiga periférica de baixa frequência	19
2.2.2 Fadiga periférica de alta intensidade	20
2.3 Fadiga central	21
2.4 Mecanismos da fadiga muscular no futebol de alta competição...	21
2.5 Monitorização de níveis de fadiga em atletas de alta competição.	24
2.5.1 Escalas de percepção de esforço	26
2.5.2 Estudo da função neuromuscular.....	27
2.5.3 Estudo da variabilidade da frequência cardíaca	28
2.5.4 Estudo da flexibilidade das articulações.....	29
2.5.5 Estudo bioquímico, hormonal e imunológico	30
2.6 <i>Data mining</i>	32
2.6.1 Descrição dos modelos de caracterização e discriminação.....	36

2.6.2	Descrição dos modelos de padrões frequentes, associação e correlação	37
2.6.3	Classificação e regressão para análise preditiva	38
2.6.4	Análise de <i>clusters</i> e <i>outliers</i>	40
3.	Descrição do <i>dataset</i> pré processamento	42
4.	Metodologia	50
4.1	Metodologia de aquisição de dados	51
4.2	Métodos de processamento do <i>dataset</i>	54
4.3	Descrição da população de estudo	57
4.4	Metodologia de análise	58
5.	Resultados	60
6.	Discussão	64
6.1	Trabalho Futuro	66
	Referências	67

Acrónimos

NFL – National Football League

NBA – National Basketball League

FCF – Futebol Clube de Famalicão

GPS – Sistema de Posicionamento Global

CMV – Contração Máxima Voluntária

ATP – Adenosina Trifosfato

DALDA – Daily Analysis of the Life Demands for Athletes

SJ – Squat Jump

CMJ – Counter Movement Jump

BPL – British Premier League

VFC – Variabilidade de Frequência Cardíaca

CK – Creatinaquinase

OLAP – Online Analytical Processing

DC – Defesas Centrais

GR – Guarda Redes

DL – Defesas Laterais

MC – Médios Centrais

EX – Extremos Laterais

AV – Avançados

RPE – Rate of Perceived Exertion

MD+1 – Match Day + 1

MD-2 – Match Day - 2

Índice de figuras

Figura 2: Variação do torque máximo em idosos e jovens.....	16
Figura 3: Comparação de tempo até falha entre idosos e jovens	17
Figura 4: Comparação do tempo até falha entre homens e mulheres.....	18
Figura 5: Mecanismos de contração e excitação e utilização de ATP celular	19
Figura 6: Futebolista a pontapear uma bola.....	24
Figura 7: Exemplo de questionário de <i>wellness</i> de Martin Buchheit.....	27
Figura 8: Exemplo de CMJ	28
Figura 9: Atleta utilizando cardiofrequencímetro <i>FirstBeat</i>	29
Figura 10: Avaliação da flexibilidade da articulação do joelho.....	30
Figura 11: Etapas para extração de conhecimento a partir de dados	33
Figura 12: Exemplo de base de dados relacional.....	34
Figura 13: Modelação de uma <i>datawarehouse</i>	35
Figura 14: Modelação de cubo OLAP	35
Figura 15: Árvore de decisão, regras <i>If-Then</i> , rede neuronal	39
Figura 16: Atleta do FCF utilizando o sensor de GPS Johan Sports.....	51
Figura 17: Avaliação <i>Squeeze test</i> 45°.....	52
Figura 18: Avaliação isométrica de extensão do joelho a 30°.....	53
Figura 19: Avaliação de exercício de iquiotibais Nórdico.....	53
Figura 20: Distribuição dos valores de <i>Squeeze</i> 45°	55
Figura 21: Modelo relacional do <i>dataset</i>	59
Figura 22: Boxplot de caracterização de força exercida em <i>Squeeze</i> 45° em MD-2 e MD+1	60
Figura 23: Gráfico linear de progressão dos níveis de força em <i>Squeeze</i> 45° durante a época 2018/2019	61
Figura 24: Boxplot de caracterização das ações de aceleração em jogo e treino.	62
Figura 25: Boxplot de caracterização das ações de <i>High Sprint</i> em jogo e treino	62
Figura 26: Boxplot de caracterização da distância total percorrida em situações de jogo e treino	63

Índice de tabelas

Tabela 1: Exemplo de dados transacionais.....	36
Tabela 2: Tabela <i>load</i> pré processamento.....	47
Tabela 3: Tabela <i>force</i> pré processamento	49
Tabela 4: Tabela <i>force</i> pós processamento.....	54
Tabela 5: Análise estatística da tabela <i>load</i>	56

1. Introdução

Motivação: Em contexto desportivo de alta competição é de grande relevância a monitorização e acompanhamento dos níveis de fadiga muscular para uma otimização da performance e idealmente, prevenção de lesão.

Sumariamente pode-se inferir que, a fadiga muscular tem como consequência a incapacidade de produção de força máxima.

O sprint, o salto e o remate são de extrema importância no sucesso de um atleta de futebol e têm relação direta com a produção de força máxima nos vários grupos musculares dos membros inferiores.

Um atleta lesionado significa para o clube perdas económicas, muitas vezes significativas, para a instituição e atrasa o crescimento e desenvolvimento do atleta. Esta não é apenas uma constatação que se pode perceber pelo estado atual dos conhecimentos na área do treino, mas vem, inclusive, sendo um dos pontos de interesse quer dos próprios clubes quer da comunidade científica e tecnológica que a cada dia desenvolve novas e melhores ferramentas para a gestão de fadiga e prevenção da lesão.

Um estudo durante três épocas consecutivas com uma equipa de elite ^[1] escreve: O primeiro passo para a otimização do rendimento de um desportista é que este se encontre em condições de treinar continuamente, neste sentido, ter os melhores jogadores permanentemente disponíveis e em boas condições para competir e treinar pode ser um fator decisivo para o sucesso desportivo.

Deste modo, a prevenção de lesões deverá ser encarada como um dos objetivos primários para o rendimento desportivo de excelência.

Hägglund et al.^[2] publica em 2013, um estudo onde analisa 11 épocas de 24 equipas de futebol profissional de 9 países europeus, perfazendo 155 épocas individuais, tanto a nível nacional como ao nível das competições europeias, foram reportadas 7792 lesões durante 1 026 104 horas de exposição, o que totaliza uma incidência de 7,7 lesões/1000 horas, sendo que em média cada equipa teve 52 lesões por época, o que equivale aproximadamente a 2 lesões por cada jogador por época. Encontraram associações significativas entre baixas

taxas de lesões e melhores desempenhos tanto a nível doméstico como europeu. (Ekstrand et al.)^[3] publica em 2011 um estudo sobre a incidência lesiva no futebol profissional, analisaram 23 equipas ao longo de 7 épocas, registaram um total de 4483 lesões ocorridas em 566000 h de exposição, dando uma incidência de 8 lesões/1000 h, os clubes pagaram cerca de 177 milhões de libras a jogadores lesionados, num total de 713 lesões.

Como não podia deixar de ser, este não é um problema apenas no futebol. Analisando as ligas de maior expressão nos Estados Unidos da América, um estudo realizado pelo Kitman Labs^[4] demonstra o impacto que as lesões têm no desempenho desportivo das equipas da NFL (principal liga de Futebol Americano) e da NBA (principal liga de Basquetebol do mundo). Os resultados obtidos após a análise de 5 épocas, mostram que, na NFL, as equipas em que a maioria do investimento salarial é canalizado para atletas saudáveis e aptos a competir simplesmente ganham mais jogos, e na NBA quem gasta mais com os jogadores não lesionados tem uma probabilidade bastante mais alta de atingir os *playoffs*.

Neste mesmo estudo demonstraram que na NFL gastam em média, por época com jogadores lesionados 10,7 milhões de dólares.

Tal como foi analisado no ponto anterior, os estudos relativamente aos custos associados às lesões e suas consequências para o atleta individualmente e/ou equipa, são cada vez mais abundantes e demonstrativos da importância do investimento nesta área. Sendo assim, sem dúvida que cada vez mais as lesões são uma das principais preocupações dos clubes, bem como dos atletas individualmente.

A busca por estratégias que possam ajudar a reduzir as lesões e suas consequências é então, uma prioridade no desporto de elite, sendo cada vez maior o investimento em estratégias/tecnologias não só para a otimização da performance mas também para gestão de fadiga e consequente prevenção de lesões.

Como tal, pretende-se neste estudo integrar e relacionar valores de testes isométricos e dinâmicos padronizados nos vários grupos musculares dos membros inferiores. Com estes equipamentos procura-se apurar os níveis de força máxima e desequilíbrio entre os membros nos grupos musculares de interesse.

No caso do equipamento Acceleration Leg Curl/Extension e Smart Groin Trainer a captação dos dados é feita através do processamento de sinais obtidos por células de carga monoaxiais. No fim do teste é exportado um ficheiro .xlsx contendo a *raw data* a ser analisada e integrada.

Objetivos: Definição de protocolos e métodos para a recolha e tratamento dos dados de forma sistemática e consecutiva durante a época futebolística de 2018/2019.

Implementação de mecanismos de data *mining* e extração de conhecimento de dados, procurando padrões relevantes relacionados com a definição de fadiga, apoiando assim a gestão dos seus níveis em atletas de futebol profissional.

Contribuição: No início da época desportiva 2018/2019, com a colaboração do coorientador Professor Carlos Pacheco, foi definido o processo e protocolo de recolha de dados.

Para os diferentes grupos musculares e equipamentos, o protocolo de recolha foi o seguinte:

- Aquisição da força máxima exercida pelo grupo muscular adutor, através de *squeeze test* 45° com o equipamento Smart Groin Trainer, Neuroexcellence.
- Aquisição da força máxima exercida pelo grupo muscular Ísquiotibial, através de teste de contração voluntária máxima de flexão do joelho a 30° em posição ventral com o equipamento Acceleration Leg Curl/Extension, Neuroexcellence.
- Aquisição do torque máximo exercido pelo grupo muscular Ísquiotibial, em exercício de extensão nórdica com o equipamento Acceleration Leg Curl/Extension, Nordic Hamstring Addapter, Neuroexcellence.

Durante o mesmo período foi também efetuada a recolha dos dados relativos à movimentação dos atletas dentro do campo utilizando tecnologia de GPS.

Estrutura: Durante época desportiva 2018/2019, os atletas selecionados do plantel principal sénior do F. C. Famalicão realizaram diferentes medições isométricas e dinâmicas dos vários grupos musculares dos membros inferiores, com a seguinte programática: 48h antes do jogo (ou treino com carga equivalente para jogadores não utilizados) e 24h após o jogo. Simultaneamente, durante os períodos de treino e jogo foi também estudado o movimento dos atletas através de sensores de GPS. Todas as sessões e jogos foram recolhidas.

Após tratamento dos dados e padronização dos mesmos, foi construída uma base de dados integrando todos os valores. Através de métodos de *data mining* e *clustering* os dados foram processados de modo à construção de um relatório interativo e dinâmico, para apoio à gestão dos níveis de fadiga dos atletas. No final da época, foi construído um relatório final, utilizando a base de dados

desenvolvida durante a mesma, de modo a apurar quais os fatores que mais contribuem para a fadiga muscular.

2. Estado da arte

Procura-se neste capítulo, analisar e descrever os mais relevantes processos inerentes à monitorização dos atletas e consequente rendimento desportivo. Terá por base a reflexão sobre as metodologias utilizadas, os diferentes eixos de análise e a qualidade, fiabilidade e relevância dos resultados obtidos.

2.1 Fadiga muscular

É conhecido que, a fadiga pode ser causada por um diferente número de mecanismos, acumulação de metabolitos dentro das fibras musculares e estímulos motores inadequado por parte do córtex motor são os principais fenómenos a ter em conta, mas não são os únicos. É então possível inferir que não existe um ponto de falha global responsável pela fadiga muscular, no entanto, este fenómeno encontra-se intimamente associada à produção sub-maximal de força nas contrações musculares.^[5]

Apesar de não ser difícil perceber quando um indivíduo se encontra fatigado, é uma tarefa completamente diferente ser capaz de identificar o mecanismo fisiológico responsável.



Figura 1: Atleta fatigado

Uma abordagem tão ampla desta temática é problemática, pois, como descrito no paragrafo anterior, a fadiga pode ser enquadrada em diferentes fenómenos e todos estes com diferentes mecanismos biológicos, resultando assim numa probabilidade reduzida de identificação concreta da causa da fadiga muscular.

Um dos protocolos mais comumente utilizados para a quantificação de fadiga é mensurar a força de uma contração máxima voluntária imediatamente antes e após de contrações fatigantes, de modo a estimar o défice relativo de força.

Avaliando a etiologia da fadiga, a abordagem mais simples seria caracterizá-la como o processo ou processos fisiológicos que inviabilizam a capacidade das proteínas contráteis de gerar força, sem descurar o local comprometido que varia conforme a tarefa a ser realizada.

Esta caracterização é conhecida por “fadiga muscular dependente das tarefas realizadas” e é um dos princípios mais antigos no conceito de fadiga.^{[6][7][8]} De acordo com este princípio, não existe uma causa única para o fenómeno e o mecanismo causador dominante está diretamente relacionado com a atividade e músculos estimulados, durante o exercício.

Atualmente não existem respostas concretas para questões como: serão os idosos mais suscetíveis à fadiga que os jovens? São as mulheres menos suscetíveis que os homens? É o estímulo neuronal que chega ao músculo é o adequado? Esta dificuldade torna-se óbvia quando analisados os seguintes estudos.

Em 2007, (Baudry et al 2007)^[9] investiga sobre a influência da idade na fatigabilidade. Para isso, avaliou o declínio do torque máximo em população adulta jovem (média $\pm$ dp.; 30.5 ± 2.5 anos) e idosos (77.2 ± 1.4 anos), homens e mulheres executaram contrações máximas de encurtamento e distensão nos dorsiflexores do tornozelo. O protocolo consistia em cinco séries de 30 contrações máximas com uma cadência de 3,5 segundos, cada repetição teve uma amplitude de 30° e a velocidade da mesma era controlada por um motor a $50^\circ/\text{s}$. Entre cada série foi definida uma pausa de 60s e as medições de encurtamento e distensão foram realizadas em dias diferentes. Os jovens adultos foram mais fortes na contração máxima como indicado pelo torque máximo registado (38.3 ± 3.1 e 28.6 ± 1.3 Nm, respetivamente).

O declínio no torque máximo foi também mais acentuado no grupo dos idosos para ambas as tarefas.

Na figura 1, podemos constatar a diminuição de 27,1% no torque máximo após 150 repetições no grupo dos jovens adultos e uma diminuição de 50,2% no grupo dos idosos. A fadiga registada em ambos os grupos foi então associada a alterações no controlo da excitação/contração de Ca^{2+} e, somente no grupo dos idosos, diminuição da propagação do estímulo neuro muscular. Estes resultados indicam então que os idosos seriam mais suscetíveis a fadiga do que jovens adultos em ações de contração e distensão dos dorsiflexores do tornozelo.

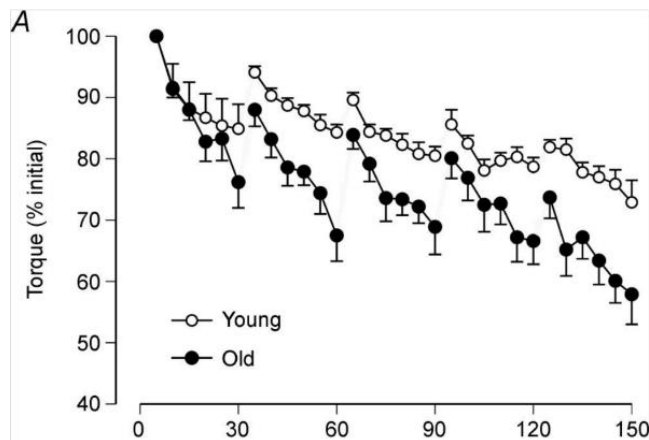


Figura 1: Variação do torque máximo em idosos e jovens

Contrastando com estes resultados, (Hunter et al, 2005)^[10] publica que indivíduos mais velhos (71.3 ± 2.9 anos) mantêm uma contração isométrica submáxima dos músculos flexores do cotovelo durante mais tempo do que jovens adultos (21.5 ± 4.4 anos). O teste consistia em manter o torque muscular em cerca de 20% da contração máxima voluntária durante o máximo de tempo possível e terminava quando o declínio no torque registado fosse igual ou superior a 10% consecutivamente durante um período temporal de 5s.

Os valores de CMV registados para o grupo de jovens adultos e idosos foram (65.9 ± 8.0 Nm; 65.4 ± 8.7 Nm, respetivamente), como é possível observar, os níveis de força de ambos os grupos são semelhantes. No entanto, o grupo de idosos registou um tempo de tensão consideravelmente superior aos jovens adultos (22.6 ± 7.4 min e 13.0 ± 5.2 min, respetivamente), como indicado na figura 2.

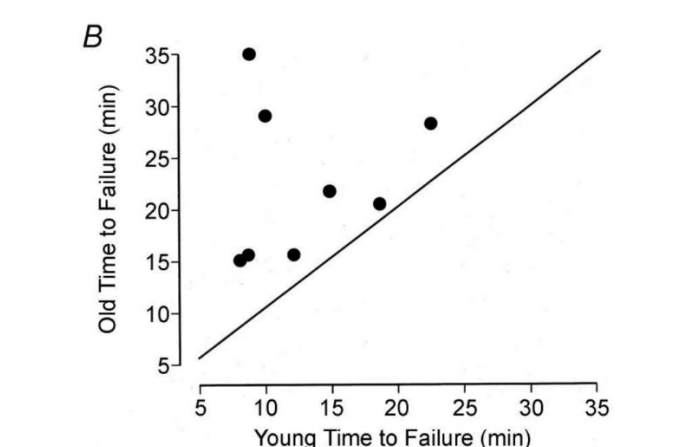


Figura 2: Comparação de tempo até falha entre idosos e jovens

O tempo até falha mais curto registado nos jovens adultos coincidiu com: aumentos mais repentinos na amplitude e frequência de atividade captada pela Eletromiografia de superfície, aumento dos níveis de percepção de esforço e aumento da média da pressão arterial e frequência cardíaca.

Estes resultados contrastam diretamente com a hipótese proposta por (Baudry et al 2007)^[9].

Descritas algumas incoerências na literatura quanto ao impacto da idade na fadiga muscular, repito o mesmo processo relativamente ao sexo.

Segundo (Hicks et al. 2001 e Clark et al. 2005)^{[11][12]} indivíduos do sexo feminino são capazes de sustentar contrações submáximas durante maiores períodos de tempo que indivíduos do sexo masculino, especialmente em contrações de baixa intensidade. Já em contrações máximas os homens suportam mais tempo que as mulheres.^[9]

A hipótese mais suportada para estes acontecimentos é que, no caso dos indivíduos do sexo masculino, é ativada maior percentagem de massa muscular e, em indivíduos do sexo feminino, há uma dependência menor de metabolismo glicolítico.

Tipicamente, os homens são mais fortes que as mulheres e, como tal, ativam uma maior massa muscular para exercer as mesmas forças relativas que as mulheres. Consequentemente existe também uma maior pressão intramuscular e maior oclusão do fluxo sanguíneo nos indivíduos do sexo masculino.^[13]

(Hunter et al., 2004)^[14] compara a fadiga entre homens e mulheres com níveis de força equivalentes e conclui que, as mulheres são de facto menos

fatigáveis em contrações submáximas intermitentes que os homens. Os dois grupos registavam valores semelhantes de torque dos flexores do cotovelo e efetivaram o teste a 20% da CMV (homens: 12.6 ± 1.6 Nm; mulheres: 12.8 ± 1.6 Nm). Não foram encontradas diferenças significativas no tempo até falha nos dois grupos (819 ± 306 e 864 ± 391 s, respetivamente). Todavia, como afirmado no parágrafo anterior, é em contrações intermitentes que surgem diferenças significativas, o grupo do sexo feminino foi capaz de manter contrações intermitentes durante um maior período de tempo, 513s para as mulheres 194s para os homens.

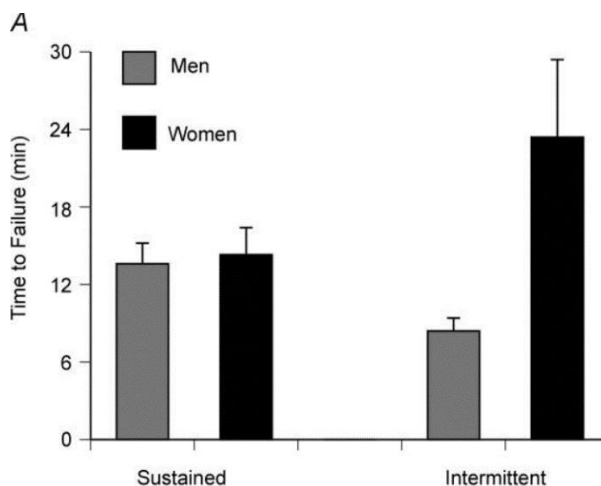


Figura 3: Comparação do tempo até falha entre homens e mulheres

2.2 Fadiga periférica

Após exercício físico, o músculo desenvolve adaptações em resposta ao tipo de treino, carga e protocolo utilizado. Independentemente destas variáveis, a intensidade de trabalho e o descanso pós treino adequado desempenham um papel fundamental no crescimento de massa muscular.[17]

No pós treino desencadeiam-se perturbações intramusculares, como por exemplo a diminuição de, glicogénio intramuscular, fosfocreatina, reservas de ATP e aumento de fosfato inorgânico e iões de hidrogénio.[18][19]

Estas perturbações tornam-se evidentes quando a intensidade da sessão de treino se aproxima ou alcança a capacidade de repetição máxima.[20]Será então pertinente quantificar e perceber quais serão os principais fatores causadores da

fadiga periférica, tal como anteriormente descrito, existem vários mecanismos fisiológicos causadores destes fenómenos com uma possível sub divisão sob as nomenclaturas que se encontram abaixo descritas.

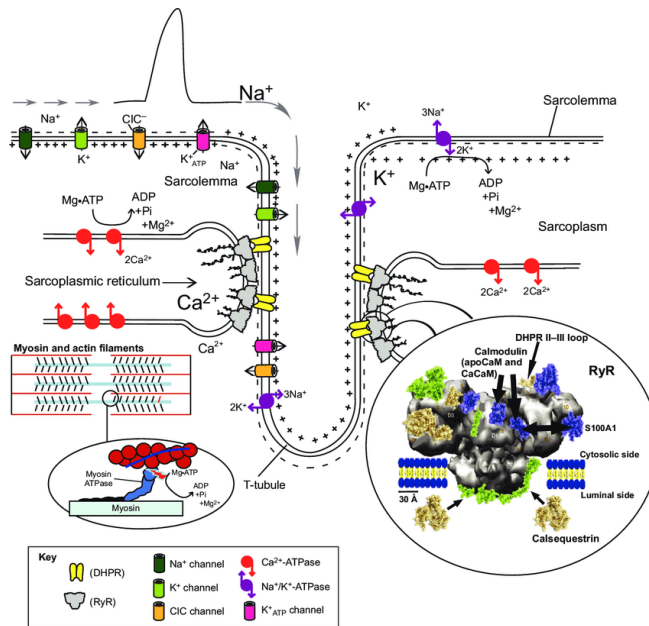


Figura 4: Mecanismos de contração e excitação e utilização de ATP celular

2.2.1 Fadiga periférica de baixa frequência

Caracterizada pela diminuição agravada da força produzida pelas fibras musculares, maior tempo de recuperação dos níveis de força e maior durabilidade de fatores de fadiga. Aparentemente, a sua recuperação relaciona-se com o débito proteico necessário para a regeneração das estruturas maculares lesadas no exercício.^{[21][22]}

As suas manifestações prolongam-se tipicamente durante várias horas ou mesmo dias.

Vários estudos sugerem que existe uma grande relação entre este fenómeno com a perda de homeostasia celular do ião Ca^{2+} , verificando duas componentes fundamentais:

- Componente metabólica, sendo esta aliviada na presença de glucose nas primeiras horas de recuperação

- Elevação dos níveis intracelulares de Ca^{2+} , cujo mecanismo de recuperação é mais moroso.

Dada a maior dificuldade de recaptção do ião Ca^{2+} , após atividade física em momento de repouso, é expectável que as concentrações intracelulares de Ca^{2+} se elevem induzindo alterações no reticulo sarcoplasmático. Foi também verificado que a acumulação deste metabolito, estimula atividade de enzimas proteolíticas e a fosfolipase A2 causadoras de degradação de proteínas e fosfolípidos da membrana. Induz o *swelling* mitocondrial e contribui para a quebra de túbulos T e reticulo sarcoplasmático. Simultaneamente, os longos períodos de isquemia, recorrentes no exercício físico, estimulam a sob produção de espécies reativas ao oxigénio, comumente associadas a mecanismos lesivos para a estrutura muscular esquelética.^{[23][24]} Em 1997, Essing e Nosek^[25] publicam um estudo onde referem que o stress oxidativo é um fator de grande impacto para a incapacidade de produção de força pelas fibras musculares.

2.2.2 Fadiga periférica de alta intensidade

Identifica-se pela diminuição de: produção de força, amplitude e duração do potencial de ação muscular agravada pelo aumento da concentração de Na^{+} intracelular e K^{+} extracelular, estando a recuperação física intimamente relacionada com o restabelecimento da homeostasia.^[26] O aumento dos níveis intersticiais de K^{+} que ocorre pelo movimento do ião para o exterior da célula durante o potencial de ação, é referenciado por vários autores como um importante fator na fadiga em exercício intenso de curta duração.^{[27][28]}

Com a falência das bombas da membrana de $\text{Na}^{+}/\text{K}^{+}$ o gradiente iónico torna-se comprometido, traduzindo-se numa consequente diminuição da amplitude do potencial de ação e da libertação de Ca^{2+} para o espaço citoplasmático. Uma consequência mais gravosa do comprometimento das bombas $\text{Na}^{+}/\text{K}^{+}$, é que com o aumento dos níveis de K^{+} intersticiais dá-se um grande desequilíbrio do potencial eléctrico das fibras musculares, e quanto maior a concentração maior o compromisso da tensão.

A par deste aumento de K^{+} está também associado um desregulamento no influxo de Na^{+} e água para o interior das fibras, agravando o compromisso dos processos de despolarização da membrana, prolongando o potencial de ação e diminuindo a libertação de Ca^{2+} para o citoplasma.^[26]

2.3 Fadiga central

O sistema nervoso é responsável pelo controlo das contrações musculares determinando quais os músculos contraídos, o número, a sequência e a força da contração. Parte do processo do treino de força é capacitar o indivíduo de produzir estímulos a um ritmo sustentável de modo a que o músculo contraia com a maior força possível.^{[29][30]}

A fadiga central pode ser descrita pela falha voluntária ou involuntária na condução do impulso elétrico no sistema nervoso promovendo assim uma menor ativação de unidades motoras e diminuição da cadência de disparo dos neurónios motores.

O impacto do sistema nervoso central na fadiga é frequentemente estudado através técnicas de contração interpolada.^[31] Para isso, é comparada a contração voluntária máxima do indivíduo com a força supramaximal produzida através de electroestimulação do nervo motor do músculo. Estudos demonstram que aparentemente, existe um mecanismo de *feedback* sensorial que inibe a descarga dos neurónios motores em condições de fadiga resultante de *feedback* reflexo dos mecanorreceptores, aparentemente sensíveis à acumulação de metabolitos.^{[32][33]}

Outros investigadores apontam para a relação entre o tempo de exercício até falha e a variação da síntese e libertação cerebral de neurotransmissores. De maior relevo aparentam ser os mecanismos de serotonina/dopamina, a cafeína como bloqueador dos recetores de adenosina em exercícios de longa duração e o impacto de aminoácidos de cadeia ramificada como inibidores da síntese de serotonina ou no aumento das concentrações de amónia plasmática causando toxicidade cerebral.

Em outros estudos é analisado o papel da acetilcolina no processo de produção de força. Este neurotransmissor é regulado pela disponibilidade de colina sérica, sendo que reduções plasmáticas de colina aparentam também ter impacto no início de fadiga.^{[34][35]}

2.4 Mecanismos da fadiga muscular no futebol de alta competição

Nos desportos de alta competição e mais concretamente no âmbito desta dissertação, no futebol, é de fundamental relevância a manutenção do performance do atleta durante um período de tempo prolongado, a fadiga é

representada pela incapacidade de manutenção dos níveis de performance esperados. Apesar deste fenómeno estar bem estudado e documentado em desportos individuais como a corrida de longa distância e ciclismo, existe comparativamente pouca atenção por parte da comunidade científica no estudo deste fenómeno nos desportos coletivos e a sua relação não só com a intensidade do exercício, mas também com as dinâmicas intra-equipa.

Como consequência da fadiga acumulada durante os jogos verifica-se um registo de declínio na força muscular que poderá estar relacionado com o aumento da predisposição do atleta para lesões nos membros inferiores. Atualmente as equipas técnicas tendem a combater este facto com um planeamento rigoroso de cargas em momentos de treino, dietas rigorosas e apropriadas, protocolos de aquecimento e utilização das substituições em dinâmicas de jogo. Com a adoção destas estratégias preventivas começam a emergir padrões que demonstram níveis de fadiga menos pronunciados, nos momentos finais dos jogos, do que em décadas anteriores onde estes protocolos e planeamentos não eram tão comuns.^[36] No entanto há ainda uma carência de investigação que confirme este padrão.

Relativamente a padrões já mais estudados, existe evidência que suporta que as depleções das reservas de glicogénio nos músculos da coxa conduzem ao declínio de performance. Foram realizadas biópsias musculares em dois grupos distintos de atletas antes, durante e após o jogo. Um dos grupos descansou no dia anterior à biópsia e o outro realizou um treino de alta intensidade. Antes do jogo os atletas que descansaram apresentaram valores de reserva de glicogénio mais de duas vezes superiores aos atletas que treinaram. Durante o intervalo os atletas que treinaram apresentavam valores residuais de reservas e os que não treinaram apresentavam já valores baixos. Como esperado e acompanhando a tendência do ensaio, no final do jogo os indivíduos treinados registaram valores quase nulos enquanto que os não treinados apresentaram valores residuais. A nível de performance, não foram registadas diferenças significativas na primeira parte, no entanto a diferença na segunda parte foi drástica onde os atletas treinados realizaram consideravelmente menos acelerações e correram menos sem bola que o outro grupo.^[37]

Outro estudo evidencia a preponderância dos níveis de metabolitos musculares e sanguíneos durante um jogo de futebol e a sua relação com ações de sprint.^[38] Os autores acompanharam 31 atletas dinamarqueses de 4º divisão registando os valores em 3 jogos amigáveis. Durante os 3 jogos registaram também os tempos de sprint em 30 metros, antes, durante e após os jogos e, como expectável, os tempos de sprint aumentaram conforme o aumento do

tempo de jogo. Foi concluído que não existia correlação entre o declínio na capacidade de sprint e o lactato muscular, pH ou conteúdo total de glicogénio e que os níveis de lactato sanguíneo não são bons indicadores do lactato muscular durante o jogo. No entanto, níveis baixos de reservas de glicogénio em fibras musculares individuais aparentam ter relação com a diminuição da capacidade de sprint.

Adicionalmente, outros investigadores procuram relações entre a temperatura corporal e o surgimento de fadiga. O aumento da temperatura para lá do valor ótimo de funcionamento parece estar relacionado com este sintoma, sendo esta situação agravada quando a temperatura ambiental é também elevada. Segundo Ekblom 1986,^[39] a distância de alta intensidade percorrida pelos atletas em situações de jogo é de apenas 500m quando a temperatura ambiental ronda os 30°C, ao passo que, em situações em que a temperatura ronda os 20°C a distância aumenta para quase o dobro, 900m. Uma das hipóteses propostas para este fenómeno é que os atletas intuitivamente diminuem a sua performance para evitar desidratação severa.

Ainda relativamente à temperatura corporal, Cheung e Sleivert 2004,^[40] afirmam que a hipertermia pode indiretamente provocar fadiga pelo aumento da tensão cardiovascular que por sua vez conduz a um défice do fluxo sanguíneo em áreas críticas ou atuar diretamente no sistema nervoso central diminuindo a capacidade de estimulação e ativação muscular.

Mohr et al., 2004,^[41] defende que o aumento da temperatura em determinados tecidos, como por exemplo os tecidos musculares, inicialmente tem um efeito benéfico na performance. Após um aquecimento de 7 minutos pré jogo, os atletas tiveram um nível de performance mais elevado na primeira parte da partida quando comparada com a segunda parte onde não fizeram nenhum tipo de aquecimento durante o período de 15 minutos de intervalo.

Concretamente no caso do futebol, os grupos musculares Ísquiotibiais e quadríceps são de extrema importância para o desempenho dos movimentos básicos da modalidade, como esperado, são também dos grupos musculares com maior incidência de lesão. Rahnema et al., 2003,^[42] utilizou um dinamómetro isocínético para captar o torque máximo destes grupos musculares, antes, após 45 minutos e após 90 minutos de exercício. O protocolo de execução foi desenhado para mimetizar o esforço e desgaste esperados numa partida de futebol. A diminuição do torque máximo, tanto em movimento concêntrico como excêntrico, foi constante com o tempo de exercício, sendo que os valores após os 90 minutos foram os mais baixos e a 45 minutos intermédios.

É esperado que todos estes fenómenos tenham consequências nos aspetos técnicos de desempenho dos jogadores. Lees and Davies 1988,^[43] avaliaram o impacto da fadiga no remate da bola. Após indução de stress muscular no atleta, foi registado um aumento na descoordenação dos segmentos da coxa e da canela resultando num remate menos veloz. Esta descoordenação afetou a transferência de energia entre a coxa e canela, que por sua vez se traduziu numa posição de impacto deficiente entre o pé e a bola.

Correntemente existem várias soluções no mercado que procuram assistir os atletas e as suas equipas técnicas na gestão e monitorização da fadiga. No seguinte tópico abordo algumas dessas soluções.

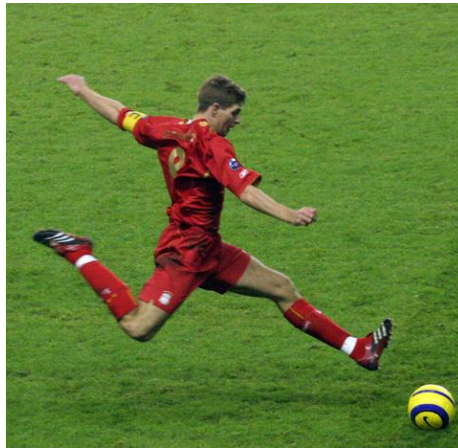


Figura 5: Futebolista a pontapear uma bola

2.5 Monitorização de níveis de fadiga em atletas de alta competição

Com o decorrer do tempo, o desporto de alta competição tem vindo a adquirir cada vez mais relevância na sociedade, não só pelo seu impacto cultural e social mas também pelo grande impacto económico. Este aumento traduz-se também numa maior procura por parte das equipas em melhorar as suas condições e práticas na busca de uma vantagem competitiva. Este crescente impacto naturalmente captou o interesse da comunidade científica de modo a desenvolverem novas e melhores ferramentas para a monitorização de um parâmetro tão relevante como a fadiga. O controlo do cansaço muscular e do estado geral do atleta poderá também oferecer informação determinante no

estudo da disponibilidade física e propensão para lesão. Métodos de avaliação rápidos, simples e não invasivos podem ser ferramentas promissoras para quantificar e estabelecer o nível de fadiga em atletas de alta competição. Uma análise robusta e precisa da flutuação dos dados recolhidos é extremamente importante para os profissionais que acompanham os atletas nas suas rotinas profissionais, no entanto, esta área ainda carece de investigação científica para apurar a sensibilidade destas ferramentas na deteção indicadores fundamentais de performance e fadiga.

Atletas de alta competição, em especial atenção nesta dissertação, futebolistas, estão expostos a grandes níveis de carga competitiva. Como consequência, a gestão dos atletas torna-se especialmente relevante pois ter todos os futebolistas disponíveis e aptos para competir aumenta significativamente a probabilidade de sucesso da equipa.^[3]

Modificações de rotina na carga de treino ocorrem durante os ciclos semanais de modo a acrescer ou diminuir os níveis de fadiga. A gestão deste fenómeno é de extrema importância para que o atleta desenvolva as adaptações musculares necessárias para ser melhor e mais forte sem comprometer a sua capacidade de competir por questões de treino excessivo.^[44] É igualmente relevante para a equipa técnica entender se o atleta se está a adaptar positivamente ou negativamente aos estímulos de treino e jogo e como é que estes afetam os níveis de cansaço muscular.

A carga de treino total reflete a carga interna e externa a que um atleta se encontra submetido. Intende-se por carga externa a capacidade de trabalho do atleta independentemente das suas limitações internas. Já por carga interna, intende-se a tensão fisiológica resultante das imposições da carga externa.

Estes conceitos têm especial impacto dado que discrepâncias entre eles podem ser diferenciadores entre atletas fatigados e não fatigados.

Apesar de serem conceitos facilmente classificados a sua quantificação em contextos desportivos de equipa pode-se tornar um processo complexo. Dada a especificidade e variabilidade dentro do desporto, seja na posição que o atleta desempenha ou no tipo de treino diferenciado para cada sector da equipa.^{[45][46][47]}

Particularmente no caso do futebol, é importante que as ferramentas para avaliação de fadiga sejam facilmente manipuláveis, pouco invasivas e de rápida administração dado que as equipas competem habitualmente uma vez por semana, sendo que em determinadas alturas da época podem chegar a jogar por três vezes na mesma semana, e como tal os atletas necessitam de estar na sua

máxima disponibilidade física com tempos de recuperação extremamente limitados.

Neste contexto as seguintes ferramentas e metodologias têm vindo a ser as mais relevantes na monitorização de fadiga em futebolistas.

2.5.1 Escalas de perceção de esforço

As escalas de perceção de esforço e questionários têm vindo a ser largamente aceites como uma ferramenta de monitorização de fadiga. Uma recente revisão bibliográfica mostra que estas metodologias obtêm maior sensibilidade para cargas crónicas e agudas de treino que medidas convencionais objetivas.^[48] Com a sua grande adoção na liga de futebol de referência a nível mundial, a BPL, foram efetuados estudos que demonstram que a perceção de esforço dos atletas estava significativamente correlacionada com as cargas de treino diárias tanto em pré-época como em períodos de competição. Da mesma maneira, questionários diários relativos à fadiga, qualidade de sono, stress e humor aparentam ter correlação com os fenómenos em cima referidos.^[49]

Algumas das escalas e questionários mais utilizados como o REST-Q e DALDA, são frequentemente alterados e utilizados como base para ferramentas próprias para os clubes pois por defeito são demasiado extensas o que desencoraja os próprios atletas a responder e os profissionais a realizar o devido tratamento dos dados.^[50]

Apesar do importante contributo que estes mecanismos de autoavaliação e perceção de esforço oferecem, não são suficientes para isoladamente servirem como sistemas de apoio à decisão da equipa técnica e clínicos na elite do desporto. Como tal, os profissionais complementam esta informação com avaliações complementares de modo a inferir sobre o estado da fadiga muscular dos seus atletas. Os métodos mais populares entre a comunidade são o estudo de: função neuromuscular, variabilidade cardíaca, flexibilidade das articulações e níveis bioquímicos, hormonais e imunológicos. Seguidamente contextualizo todas estas metodologias e qual a sua relevância no mundo do futebol.

Wellness Questions from Martin Buchheit					
	5	4	3	2	1
Fatigue	Very fresh	Fresh	Normal	More tired than Normal	Always tired
Sleep Quality	Very restfull	Good	Difficulty falling asleep	Restless sleep	Insomnia
General Muscle Soreness	Very Fresh	Feeling good	Normal	Increase in soreness/tightness	Very Sore
Stress Levels	Very Fresh	Relaxed	Normal	Feeling stressed	Highly stressed
Mood	Very positive mood	A generally good mood	Less interested in others &/or activities than usual	Snappiness at team-mates, family, and coworkers	Highly annoyed/ irritable/ down

Figura 6: Exemplo de questionário de *wellness* de Martin Buchheit

2.5.2 Estudo da função neuromuscular

O estudo da função neuromuscular através de protocolos de salto revolucionou por completo a elite do desporto pela simplicidade e rapidez com que os profissionais poderiam obter informações de relevo sobre os seus atletas. Estes protocolos de salto com o *Squat Jump* (SJ) e *Counter Movement Jump* (CMJ) foram os mais adotados demonstrando drásticas diminuições na capacidade de salto dos atletas fatigados com impacto até 72h após jogo ou estímulo semelhante.^[51]

No entanto, este método aparenta ser pouco sensível como ferramenta de avaliação diária não captando alterações nas cargas de treino. Em estudos realizados na BPL, onde os atletas realizavam diariamente estes protocolos de avaliação, foi demonstrado que estes métodos não são suficientemente sensíveis como métrica de avaliação da função neuromuscular. Complementando esta afirmação, outros estudos que avaliam a sensibilidade desta metodologia em equipas de rugby e futebol de formação demonstram que não há alterações relevantes na altura registada em CMJ ou correlações relevantes com as alterações de carga de treino. A utilização da métrica de altura de salto aparenta então ser pouco sensível para a alterações das capacidades físicas registadas no dia a dia.^[52]

Contudo, outras variáveis relativas ao CMJ como: força excêntrica, força concêntrica, duração total, tempo para máximo, potência, tempo de voo e rácio de contração/tempo aparentam apresentar maior sensibilidade para a deteção de fadiga muscular. Um estudo efetuado na liga de futebol australiano durante uma época reitera que variações nos parâmetros força tempo aparentam ser sensíveis para aumentos e diminuições na carga de treino diária.^[53]

Com o maior interesse da comunidade científica nesta temática é expectável que surjam diferentes e possivelmente melhores alternativas para o estudo da função neuromuscular em atletas de alta competição.



Figura 7: Exemplo de CMJ

2.5.3 Estudo da variabilidade da frequência cardíaca

A análise de parâmetros relativos ao tempo de resposta vagal da variabilidade da frequência cardíaca, em vez da análise mais tradicional do seu espectro, têm tomado cada vez mais o interesse da comunidade científica, dadas as superiores capacidades avaliativas conferidas por esta metodologia. De um modo geral, a VFC está reduzida nos dias imediatamente a seguir a exercício físico intenso, no entanto, estas condições são inconsistentes em atletas de desportos de endurance. Tal como no estudo da função neuromuscular, existe pouca evidência que corrobore que esta metodologia é suficientemente sensível a flutuações de cargas de treino diárias.^[54]

Em futebol de elite, existe aparente relação entre a VFC e as distâncias percorridas em corrida de alta velocidade. No entanto, um estudo na mesma população indica que a VFC não se altera significativamente durante uma semana de treino durante a época desportiva.^[55]

Curiosamente, a sensibilidade da VFC quando estabelecida a média semanal aparenta ser um bom indicador comparativo da carga semanal dada a grande variação nas medições de dia para dia, no entanto a sua aplicação em desportos coletivos pode ser desafiante com a grande necessidade de monitorização constante e processamento pós treino.^[56]

Como indicado no subcapítulo anterior, esta área também carece de uma intervenção da comunidade científica de modo a validar e melhorar estas práticas para monitorização de fadiga em atletas de alta competição.



Figura 8: Atleta utilizando cardiofrequencímetro *FirstBeat*

2.5.4 Estudo da flexibilidade das articulações

A simples avaliação clínica da flexibilidade das articulações e o seu alcance de movimento é uma avaliação rotineira de pré-época por parte das equipas técnicas no desporto de elite. Contudo existe uma grande lacuna nos estudos utilizando esta época para avaliação da fadiga em resposta a estímulos de treino e jogo. Esta avaliação mais regular pode fornecer dados de extrema relevância relativamente à fadiga e risco potencial de lesão como Ispirlidis I, et al.^[57] demonstra em 2008 com um estudo que indica a redução de flexibilidade da articulação do joelho até 48h após jogo. Mohr et al 2016,^[58] escreve também que tipicamente, no período de 24h a 48h após jogo, a flexibilidade do joelho

está comprometida em 7%. Outro estudo conduzido em futebol de formação mostrou bons resultados e validade na análise da extensão da anca, neste estudo foi registado uma diminuição da capacidade e extensiva da articulação na ordem dos 12,5%, 14h após jogo. A natureza simples desta metodologia e a capacidade de avaliar estruturas chave para o desporto em questão poderá traduzir-se num melhor conhecimento da fadiga da estrutura e o seu risco potencial de lesão associado.



Figura 9: Avaliação da flexibilidade da articulação do joelho

2.5.5 Estudo bioquímico, hormonal e imunológico

Provavelmente a área de estudo com mais suporte na comunidade científica de todos os métodos acima descritos, no entanto até à data, ainda não foi encontrado um biomarcador com resultados conclusivos para a definição de fadiga muscular em atletas de alta competição. Apesar de ser uma área altamente estudada, estas metodologias têm vindo a cair em desuso em equipas de elite dado os elevados custos e o tempo necessário para consumir a análise, tornando-se assim pouco práticas para ambientes de desportos coletivos onde as equipas são compostas por vários jogadores.^[59]

Evidenciando os estudos desenvolvidos nesta área, foram analisados vários marcadores passíveis de quantificar os níveis de fadiga. A creatinaquinase (CK) aumenta imediatamente após o término do jogo de futebol culminando o seu valor máximo num período entre as 24h e 48h, os valores normais são restabelecidos entre 48h e 120h após jogo.^[57] Apesar de ser claramente o

marcador mais utilizado, questões relativas ao mecanismo exato da sua atividade após exercício e a recuperação da função muscular ainda não estão respondidas. Já a proteína C Reativa e os níveis de ácido úrico demonstram ser um marcador de inflamação sensível no caso de futebol concretamente. Aumentos de sensivelmente 50% no nível de proteína C reativa 48h após o jogo foram observados em jogadores de futebol de elite, assim como, um aumento dos níveis de ácido úrico atingindo valores 72h após jogo. Outros marcadores como os níveis de cortisol, testosterona e IL-6 são frequentemente referenciados na literatura. ^[58]

Apesar da grande cobertura científica nesta área, existe uma grande falta de estudos longitudinais em desportos coletivos de elite, especialmente englobando períodos de competição e treino. Esta lacuna poderá estar muito associada ao facto dos métodos de recolha serem extremamente invasivos para os atletas, ser necessário um compromisso de tempo elevado por parte do departamento médico dos clubes e o próprio custo de manutenção de análises rotineiras por parte de todo o plantel.

2.6 *Data mining*

[60]Cada vez mais se verifica a importância e relevância dos dados na sociedade atual, provenientes de todas áreas socioeconômicas a recolha de dados está presente em todos os aspetos do nosso quotidiano. Este crescimento explosivo em volume de dados é uma clara consequência da informatização da nossa sociedade e o desenvolvimento de mais e melhores ferramentas de captação e armazenamento de dados.

A um ritmo desenfreado de crescimento os dados tomam proporções tão massivas que um novo paradigma surge. Com tanta informação torna-se humanamente impossível retirar conhecimento dos dados sem uma maior automação do processo. Com base nesta premissa, surge o *data mining*, uma área bastante jovem, mas dinâmica e promissora capaz de suportar e dinamizar a crescente necessidade de conhecimento por parte das gerações atuais e futuras.

Como é que o *data mining* transforma dados em conhecimento?

Dado o exemplo do motor de busca de referência a nível mundial, o google, este recebe diariamente biliões de *queries* em que cada uma é uma transação onde o utilizador descreve a informação que procura. Assim, alguns padrões começam a surgir, do qual se pode tirar conhecimento relevante que não seria possível constatando as *queries* individualmente.

Pode-se enquadrar o desenvolvimento da *data mining* como um processo natural evolutivo da tecnologia. Para este crescimento foi de fundamental impacto a evolução das tecnologias de base de dados e sua gestão.

O que é então o *data mining*?

Sendo um tema interdisciplinar torna-se difícil formular uma definição concreta, dado que o próprio termo não representa a maioria dos componentes envolvidos no processo. Uma designação mais completa destes processos seria “extração de conhecimento a partir de dados” que peca por ser longo e pouco prático. Alguns autores definem *data mining* com um processo em extração de conhecimento a partir de dados outros identificam as duas expressões como sinónimas.

De um modo sucinto fazem parte deste processo as seguintes etapas:

- Limpeza dos dados, onde deverá ser removido o ruído e dados inconsistentes
- Integração dos dados, onde se combinam os dados obtidos por diferentes fontes
- Seleção dos dados, onde serão selecionados os dados relevantes para a fase de análise

- Transformação dos dados, onde deverão ocorrer transformações sumárias ou de agregação de modo a que os dados sejam adequados para os diferentes algoritmos
- *Data mining*, onde serão aplicados diferentes algoritmos e/ou metodologias para extrair padrões e/ou conhecimento
- Avaliação de padrões, fase onde deverá ser feita uma avaliação dos padrões descobertos de modo a inferir sobre a sua relevância
- Apresentação de conhecimento, por fim sob um conjunto de regras e técnicas deverá ser feita a visualização do conhecimento de forma sucinta e clara.

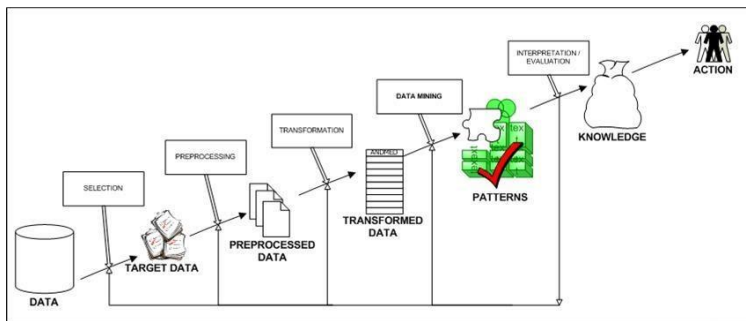


Figura 10: Etapas para extração de conhecimento a partir de dados

Os processos 1 a 4 compreendem o pré-processamento de dados, onde estes são preparados para os algoritmos a serem implementados. Uma prática comum nos dias que correm é efetuar estes passos previamente a gravação dos dados numa *data warehouse* ou outra base de dados.

Como referido anteriormente alguns autores defendem que o *data mining* é então uma pequena parte de algo muito maior (extração de conhecimento a partir de dados), sendo essa a referência utilizada na caracterização. No entanto, o padrão da indústria é a utilização do termo *data mining* e como tal será essa a nomenclatura adotada no decorrer da dissertação. *Data mining* é então o processo de descoberta de padrões e conhecimento através de grandes quantidades de dados de variadas fontes.

Qualquer tipo de dados pode ser utilizado para estes processos assumindo que estes são relevantes para o objetivo final. Exemplos comuns são os dados provenientes de bases de dados, *datawarehouse* e dados transacionais. No caso das bases de dados, estas consistem em dados inter-relacionados que podem ser consultados e alterados através de softwares específicos como por exemplo o Microsoft Excel. Uma base de dados relacional é um conjunto de duas ou mais tabelas que agregam atributos (colunas) e tipicamente grandes amostras de

linhas. A cada linha está associado uma chave única que é descrita pelos vários atributos das diferentes tabelas. O acesso a estas bases de dados pode também ser feito através de *queries* escritas em linguagem relacional (SQL), uma dada *query* é transformada num conjunto de operações com impacto direto nos processos e dados. Este método é especialmente útil para procurar *subsets* dentro da base de dados ou utilizar funções de agregação como somas, médias, contagens, maior, menor, etc.

Usando processos de *data mining* em bases de dados relacionais, é possível pesquisar por tendências, padrões e desvios da norma. A título de exemplo, em situação empresarial, é possível inferir que as vendas de determinado produto estão mais baixas quando comparadas com o período homólogo do ano passado. Com maior investigação é também possível perceber a causa desta redução.

As bases de dados relacionais são extremamente comuns e um dos repositórios de informação mais ricos sendo assim a maior fonte de dados para processos de *data mining*.

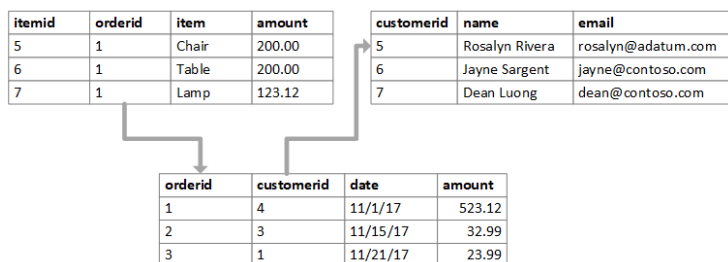


Figura 11: Exemplo de base de dados relacional

Concluída a caracterização das bases de dados relacionais introduzo o conceito de *datawarehouses*. Supondo que uma determinada empresa multinacional deseja construir relatórios de vendas de determinado produto de todas as suas lojas a nível mundial é necessária a conjugação das várias bases de dados de cada filial, o que se torna extremamente complicado se cada uma estiver guardada em servidores locais em diferentes pontos do globo. No entanto, com um sistema como *datawarehouse* é possível criar um repositório que recolhe informação de diferentes fontes e as guarda sob um *schema* único, tipicamente num só local físico. Tipicamente, de modo a facilitar a navegação, os dados nas datawarehouses estão gravados e organizados por categoria com o objetivo de oferecer conhecimento com uma perspetiva histórica.

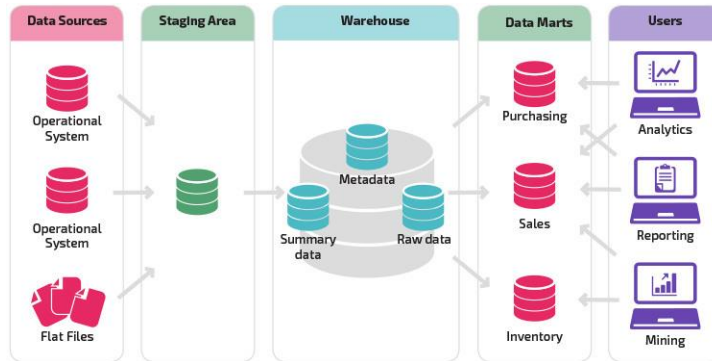


Figura 12: Modelação de uma *datawarehouse*

Esta arquitetura é normalmente modelada por uma estrutura chamada cubos de dados. Cada “face” corresponde a uma dimensão ou conjunto de dimensões definidos pelo *schema* e cada célula do cubo grava uma métrica agregada. Com visualizações multidimensionais e sumarizações pré computacionadas aglomeradas no cubo torna-se possível oferecer serviços OLAP (*online analytical processing*). Estes serviços oferecem possibilidade de extração de conhecimento em diferentes níveis de abstração definidos pelo que o utilizador procura. Apesar deste processo já suportar uma análise dos dados fornecidos, são necessárias mais ferramentas para análises mais profundas de *data mining*. Estas ferramentas e processos são conhecidos como *data mining* multidimensional.

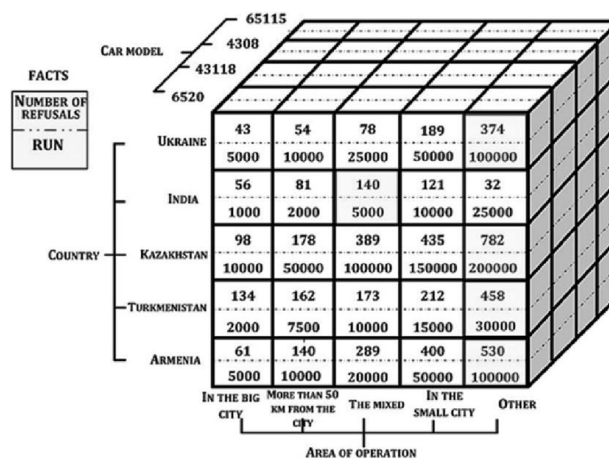


Figura 13: Modelação de cubo OLAP

Por fim, relativamente ao tipo de dados a serem utilizados para processos de *data mining*, abordo os dados transacionais. Numa base de dados transacional todas as entradas registam uma transação que tipicamente se encontra associada a um número identificativo único de transação e uma lista de itens que descrevem a atividade realizada. Esta lista de itens é normalmente descrita com tabelas de suporte que contêm informação relativa a esse acontecimento ou facto. O processo de *data mining* utilizado para dados transacionais é um modelo de padrões frequentes denominado conjunto de itens frequentes, onde se procuram padrões relativos às frequências das transações.

Tabela 1: Exemplo de dados transacionais

Transação	Item
T1	I1, I2, I3
T2	I3, I4, I10
T3	I1, I5, I10

Existem diferentes funcionalidades nos processos de *data mining*. Exemplo destes processos são: caracterização e discriminação, padrões frequentes, associações e correlações, classificações e regressões, análise de clusters e análise de *outliers*. Estas tarefas podem ser classificadas em duas grandes categorias, descritivas e preditivas. Como o próprio nome indica, os processos descritivos caracterizam propriedades encontradas na base de dados e os processos preditivos utilizam metodologias de previsão baseadas nos dados da base de dados.

2.6.1 Descrição dos modelos de caracterização e discriminação

Os dados podem ser associados a classes ou conceitos que não são mais do que descritivos de uma determinada entrada na base de dados.

Os modelos de caracterização são a sumarização das características gerais de uma determinada classe. Existem várias formas de representação de resultados e consequente visualização deste modelo, as práticas mais comuns consistem em gráficos de setores, de barras, curvas, cubos multidimensionais e tabelas

multidimensionais. Um exemplo concreto desta metodologia seria, o treinador principal do FCF deseja saber quais são os atletas que correm mais de 7 km por jogo. Os resultados seriam jogadores que ocupam posições do meio campo, têm elevados níveis de força e jogam os 90 minutos. Com este processo o treinador poderia escolher detalhar mais ou não a granularidade da resposta.

Já os modelos de discriminação são obtidos através da comparação das classes em questão com características gerais de uma ou mais classes contrastantes. As características a serem avaliadas deverão ser selecionadas através de *queries* às classes de estudo na base de dados. Os modos de apresentação deverão ser os mesmos que referidos nos modelos de caracterização com a especial atenção de que deverão estar presentes medidas de comparação visíveis que ajudam o utilizador a diferenciar as classes em estudo. A título de exemplo, o treinador principal do FCF deseja comparar os defesas com os atacantes, o resultado desta metodologia deveria referenciar métricas comparativas como, os defesas têm 40% mais força de adutores que os atacantes, os atacantes correm mais 15% que os defesas, entre outros.

2.6.2 Descrição dos modelos de padrões frequentes, associação e correlação

Tal como o nome evidencia, padrões frequentes, são padrões que ocorrem frequentemente na base de dados, atualmente classificam-se em diferentes tipos entre os quais: conjunto de itens, subsequência e subestrutura frequentes.

Os conjuntos de itens frequentes referem-se não mais do que a conjuntos de itens que tipicamente são encontrados juntos em bases de dados transacionais, as subsequências frequentes referem-se a padrões na ordem das entradas dos dados e por fim as subestruturas frequentes indicam estruturas que apresentam um padrão de relação com itens da própria estrutura ou outras subsequências.

Análise de associação é especialmente útil para a descoberta de relações potencialmente interessantes escondidas em grandes bases de dados, para isso são utilizadas regras de associação descritas pela expressão $X \rightarrow Y$, onde X e Y são conjuntos de dados diferentes.

A força da regra de associação pode ser avaliada pelo seu suporte e confiança. Uma regra com baixo suporte pode simplesmente acontecer pelo acaso já a confiança como o próprio nome indica a validade de uma inferência da regra.

- Suporte de uma regra de associação:
 - $\sigma(X)$ é a contagem de suporte de X
 - N é a contagem de transações
 - $s(X \rightarrow Y) = \sigma(X \cup Y) / N$
- Confiança de uma regra de associação:
 - $\sigma(X)$ é a contagem de suporte de X
 - $\text{conf}(X \rightarrow Y) = \sigma(X \cup Y) / \sigma(X)$

As regras de associação são frequentemente descartadas se não satisfizerem um valor mínimo de suporte e confiança, no caso destes valores serem positivos é possível inferir sobre possíveis correlações estatísticas interessantes com uma análise futura adicional.

2.6.3 Classificação e regressão para análise preditiva

Entende-se por classificação o processo de descoberta do modelo ou função que poderá descrever e distinguir as classes de dados. Estes modelos são proposições baseadas na análise de bases de dados de treino onde as classes de dados e os seus objetos são conhecidos. Após este treino o modelo deverá ser capaz de identificar uma classe desconhecida a partir de dados conhecidos, sendo que pode ser descrito por diferentes expressões e métodos. Atualmente os métodos mais utilizados são: regras de classificação (regras *IF-THEN*), árvores de decisão, fórmulas matemáticas e redes neurais.

A árvore de decisão é um fluxograma em que cada nóculo representa um teste sob determinado valor e cada linha representa o resultado desse teste, é um modelo que pode facilmente ser convertido para regras de classificação. Contrastando com a aparente simplicidade das árvores de decisão, as redes neuronais, quando utilizadas para classificação, são definidas por um conjunto de unidades de processamento semelhantes a neurónios ligadas entre si por relações com diferentes pesos, este famoso modelo é inspirado pelo sistema nervoso central dos humanos. Outros processos conhecidos de classificação são: classificação Naive Bayes, máquinas de vetor suporte e K-vizinhos mais próximos.

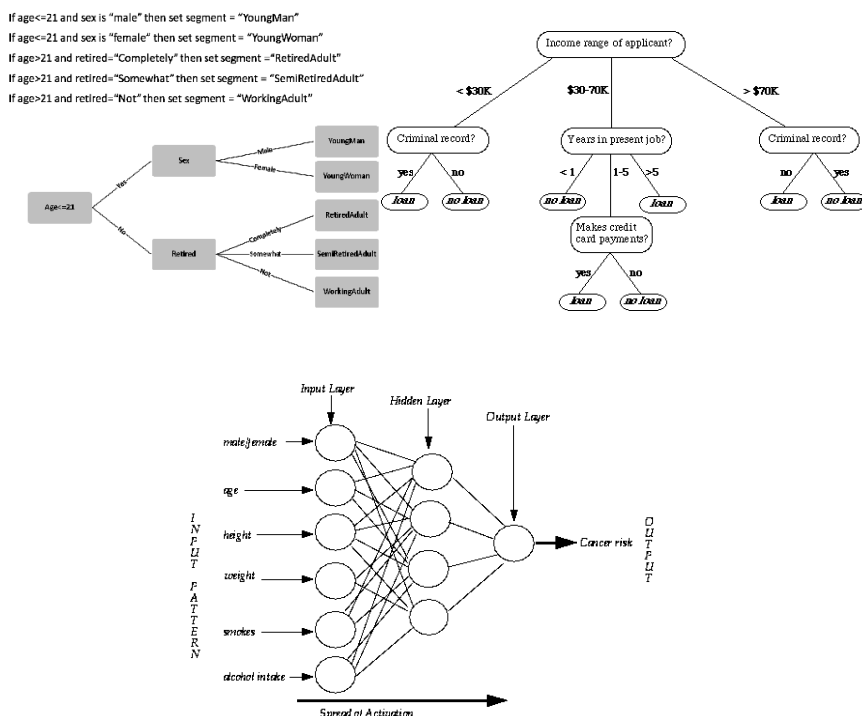


Figura 14: Árvore de decisão, regras *If-Then*, rede neuronal

Enquanto que todos estes modelos de classificação preveem classes categóricas os modelos de regressão são capazes de prever valores contínuos. Simplificando, os modelos de regressão são utilizados para previsão de dados numéricos em falta ou indisponíveis, sendo este um método analítico e estatístico capaz de avaliar a tendência da distribuição com base nos dados disponíveis. Hoje em dia, é frequentemente utilizado para estimativa de probabilidade condicional da variável dependente para dada variável independente. Exemplos conhecidos deste modelo são a regressão linear, simples, polinomial, entre outros.

Uma boa prática antes da aplicação de qualquer um destes modelos é executar uma análise de relevância onde se procura identificar possíveis

atributos que sejam mais relevantes para a classificação e/ou regressão. Estes atributos deverão ser incluídos no processo, ao passo que aqueles que forem identificados como irrelevantes não deverão ser tomados em consideração nos modelos.

2.6.4 Análise de *clusters* e *outliers*

Como descrito no sub capítulo anterior, os modelos de classificação e regressão analisam bases de dados em que os dados já se encontram organizados e classificados em diferentes classes, no entanto, em análise de *clusters* os dados ainda não se encontram classificados, muitas vezes porque simplesmente as classes ainda não foram definidas, sendo que através deste processo se torna possível classificar os dados em diferentes grupos. Estes são agrupados de modo a maximizar a similaridade intra classe e minimizar a similaridade inter classe, resumindo, as classes são formadas de modo a que os dados dentro delas sejam o mais similares possíveis mas ao mesmo tempo o mais díspares dos dados das outras classes.

Este processo possibilita também a introdução do conceito de taxonomia, onde é possível definir uma hierarquia de classes que se agrupam em eventos semelhantes.

Relativamente à análise de *outliers* esta metodologia poderá ter especial relevo em determinadas condições. É certo que várias bases de dados contêm dados que não obedecem às tendências do modelo, os chamados outliers, assim por defeito a grande maioria dos modelos de *data mining* excluem estes dados como ruído ou exceções. No entanto, há certas aplicações e contextos como por exemplo a busca de leituras anormalmente baixas de força de adução da bacia, podendo assim servir de indicador prévio para momentos lesivos. A deteção destes valores poderá ser feita através de testes estatísticos que assumem a distribuição da base de dados ou através de métricas de distância dos dados a determinados clusters. Num nível de abstração mais baixo, os modelos baseados em densidade poderão identificar *outliers* em pequenas regiões que num contexto global de distribuição estatística aparentavam enquadrar-se com os valores vizinhos.

Todos os processos acima descritos, no presente capítulo, são capazes de gerar milhares de padrões ou regras, no entanto, nem todos estes serão relevantes para a extração de conhecimento e, tipicamente, apenas uma pequena parte destes se revela interessante para o utilizador. Então o que torna

um padrão relevante? Um padrão é relevante se for facilmente entendido por humanos, é válido quando aplicado numa base de dados com algum grau de certeza e é potencialmente útil na informação que transmite. Um padrão interessante é aquele que representa e transmite conhecimento!

3. Descrição do *dataset* pré processamento

No início da época futebolística de 2018/2019, o departamento de alto rendimento do FCF propôs-se a desenvolver uma metodologia de avaliações analíticas diárias. Com estas avaliações foi possível construir uma base de dados, tema desta dissertação, que serviu de apoio à decisão e ferramenta de gestão de fadiga para a equipa técnica.

Neste *dataset* constam 2 tabelas relacionais com dados não tratados, captados pelo GPS, Johan Sports, e pelos equipamentos de avaliação de aplicação de força, Neuroexcellence.

^[61]O modelo relacional baseia-se no facto de todas as entradas de dados estarem guardadas em tabelas interligadas por um fator comum formalizador da relação. Este fator é conhecido como chave estrangeira que não é mais do que uma coluna na base de dados que estabelece a relação entre duas ou mais tabelas. Por cada valor na coluna designada como chave estrangeira, existe um valor correspondente noutra tabela. Como tal, é necessária uma especial atenção na manipulação deste campo, a inserção ou deleção de valores pode destruir a relação entre tabelas.

Após a conclusão de cada sessão de avaliação ou de treino, os dados não tratados captados pelos sensores foram introduzidos manualmente em 2 tabelas. Na tabela *load* foram manualmente inseridos todos os dados recolhidos pelo sensor de GPS, o dispositivo dispõe de uma plataforma online de onde é possível descarregar um ficheiro *Excel* (.xlsx) contendo a informação relativa à sessão de interesse. Já a tabela *force* teve como objetivo a recolha de medidas de aplicação de força em diferentes tipos de exercícios, de modo a avaliar ou estimular a produção de força em grupos musculares distintos.

A coluna “id_jogador” foi utilizada como chave estrangeira de modo a interrelacionar ambas as tabelas, por motivos de proteção de privacidade dos

atletas a identidade dos mesmos não será revelada, sendo utilizada para a sua referência nesta dissertação o atributo “id_jogador” constituído por 3 caracteres alfanuméricos.

Avanço agora para a descrição detalhada dos atributos incluídos em ambas as tabelas procurando justificar a sua inclusão ou exclusão do âmbito desta dissertação.

Numa ótica de descrição geral do *dataset* começo por uma abordagem alargada da qual posteriormente individualizo para este estudo de caso.

Como referido anteriormente as duas tabelas relacionais *load* e *force* contêm os dados não tratados captados pelos diferentes sensores, assim sendo, o tamanho da base de dados foi continuamente crescendo com o desenrolar da época desportiva.

Começando pela tabela *load* esta é constituída pelos seguintes atributos:

Data, como o próprio nome indica são dados do tipo data utilizados para uma maior facilidade de pesquisa e compreensão temporal do *dataset*, o valor mais baixo registado é 13 de agosto de 2018 e o valor mais alto, 9 de maio de 2019, tendo o ensaio a duração total de 269 dias.

Semana, são dados categóricos do tipo texto que tal como o atributo data tem o propósito de facilitar a pesquisa e associar o período de trabalho à jornada desportiva que tipicamente se disputa em intervalos de uma semana. A tabela *load* tem então a duração total de 40 semanas.

Jogador, como referenciado anteriormente neste capítulo, este atributo tem a particularidade de ser a chave estrangeira que relaciona as duas tabelas. É um atributo categórico do tipo texto do qual fazem parte 30 atletas distintos cuja identidade não será revelada no decorrer desta dissertação, no entanto, nem todos os atletas completaram a totalidade da época desportiva no plantel principal do FCF e outros integraram a equipa durante a abertura do mercado de Inverno, este acontecimento será particularmente descrito durante as metodologias, pois apenas atletas que completaram a totalidade da época desportiva serão considerados para avaliação do estudo.

Posição, são também eles dados categóricos do tipo texto, a sua organização segue a proposta de Andrzejewski, et al. 2015,^[62] em que os atletas são agrupados conforme a sua posição principal de jogo. Temos então os seguintes dados: GR (Guarda-Redes), DC (Defesas Centrais), DL (Defesas Laterais), MC (Médios Centro), EX (Extremos Laterais), AV (Avançados). Este agrupamento surge pela necessidade de agrupar atletas com estímulos de jogo equiparáveis, o que teoricamente, induziria estímulos de fadiga semelhantes e possibilita à equipa técnica análises de determinado atleta quando comparado com a média

da sua posição. Pelo seu contributo altamente específico, a posição de GR está sujeita a um desgaste totalmente diferente do que os restantes jogadores de campo, como tal, esta posição não será contemplada no decorrer da dissertação.

Peso Rec, são os primeiros dados contínuos do tipo numérico presentes da tabela *load*, registam o peso inicial dos atletas e teve como principal propósito servir de termo comparativo para as flutuações de peso durante a época. A média da equipa é de $79,96 \pm 7,69$ com valor máximo de 92 e mínimo de 63.

Peso At, são dados contínuos do tipo numérico que indicam o peso à data da medida, foram registados não só com o objetivo de observar as alterações de peso dos atletas, mas também para base de cálculo do rácio de força do grupo muscular adutor com o peso corporal. Esta métrica poderá ser um indicador precoce de lesão de adutor com descrito por Moreno-Pérez, et al. 2019.^[63] A média de equipa registada foi de $78,61 \pm 8,65$ kg. Neste atributo são registados os primeiros valores *missing* da tabela, justificados pela incapacidade do cumprimento do protocolo pela equipa técnica. O ambiente altamente volátil do contexto desportivo do futebol de alta competição dita que o acompanhamento científico nem sempre é uma prioridade, sendo que as decisões técnico táticas do treinador principal são soberanas no planeamento protocolar do dia a dia da equipa. Durante a época foi também registada uma alteração do treinador principal e alguns membros da equipa técnica, como tal, houve necessidade de discutir e reajustar os procedimentos com o novo treinador principal. A contagem de valores *missing* é de 713. Não foram realizadas quaisquer substituições nos valores *missing* nem cálculos de suavização, tendo sido apenas excluídos da análise.

Estado, contem dados categóricos do tipo texto relativos ao tipo de trabalho executado pelo atleta em determinado dia. A título de exemplo, o atleta em determinado dia poderá não ter treinado por se encontrar lesionado ou em situação normal realizou o treino de campo agendado.

RPE; *rated perceived exertion* ou em português, escala de perceção de esforço, contém dados contínuos do tipo numérico. Como referenciado no capítulo de estado da arte, estas escalas têm vindo a adquirir grande relevância ao longo dos anos quer a nível de aceitação por parte dos atletas e técnicos de equipas de alta competição quer por parte da comunidade científica que tem explorado cada vez mais esta temática. A escala é compreendida entre os valores de 1 a 10 sendo a média de resposta $6,854 \pm 1,914$. Neste atributo foram assumidos 2168 valores *missing* perfazendo aproximadamente 32% da totalidade da tabela. Estes valores poderão estar associados a períodos de folga e incumprimentos do protocolo de avaliação por parte da equipa técnica.

Total Distance, ou em português, distância total, são dados contínuos do tipo numérico adquiridos pelo sensor GPS. Como o próprio nome do atributo indica, os dados são referentes à distância total percorrida pelos atletas durante determinada sessão. Devido a alguns problemas de fiabilidade do material, foi recorrente durante a época desportiva os equipamentos perderem o sinal de comunicação entre o recetor e os sensores introduzindo assim viés de amostra. De todo o modo, a média do plantel durante a época de 2018/2019 foi de $5497,98 \pm 2325,51$ m, com o valor máximo registado de 15767,78 e valor mínimo de 20. Neste atributo foram registados 2769 valores *missing* correspondentes a 40% da totalidade. Como justificado anteriormente esta alta percentagem poderá estar relacionada com anomalias técnicas, períodos de descanso e incumprimentos na utilização do sensor.

Walk Distance, ou distância percorrida a passo, são dados contínuos do tipo numérico. Captados pelo sensor GPS estes dados registam a distância que o atleta percorre em velocidades inferiores a 7 km/h. O valor mínimo registado foi de 66,77 m e o máximo de 6808,25 m, sendo a média do plantel $2845,98 \pm 869,62$ m. Foram também registados 2791 valores *missing* contemplando 41% da totalidade.

Jog Distance, valores contínuos do tipo numérico relacionados com a distância percorrida pelo atleta em corrida lenta. Este tipo de corrida foi definido por velocidades compreendidas entre os 7 km/h e os 14 km/h. Neste atributo foi registada uma média de $1858,27 \pm 1109,71$ m com o valor máximo de 7696,22 m e 1,43 m de mínimo. Relativamente aos valores *missing* perfazem aproximadamente 41% da totalidade com o valor absoluto de 2785.

Run Distance, também entendida com a distância em corrida, são valores que registam a distância total percorrida pelo atleta em velocidades entre os 14 km/h e os 20 km/h. A média registada foi de $620,05 \pm 524,04$ m, valor máximo de 3412,07 m e valor mínimo 0,21 m. A contagem de valores *missing* é de 2829 equivalentes a aproximadamente 42%.

Sprint Distance, é também um atributo derivado dos sinais recolhidos pelo sensor de GPS que contem variáveis contínuas do tipo texto. É relativo à distância percorrida pelos atletas em momentos de *sprint*, parametrizado entre os valores de 20 km/h e 25 km/h. O valor máximo registado foi de 961,95 m e o valor mínimo de 0,19 m com uma média global de $154,38 \pm 155,55$ m. Foram também registados 2997 valores *missing* correspondentes a aproximadamente 44% da amostra.

Max Sprint Distance, os valores associados a este atributo são também eles contínuos do tipo numérico. Traduzindo para português temos a distância de

sprint máximo que está reservado para valores de velocidade superiores a 25 km/h. O valor mínimo observado foi de 0,07 m e o máximo 500 m, a média registada foi de $53,72 \pm 64,25$ m. Relativamente aos valores *missing*, neste caso concreto representam 58% da totalidade da amostra com um valor de 3980 casos. Aliado às justificações anteriores sobre as altas percentagens de valores em falta relativos ao sensor GPS é necessário também ter em conta que uma elevada percentagem de atletas não atinge estes valores de velocidade em ações de treino ou jogo.

Max Speed, como indicado pelo próprio nome, é um atributo relativo contínuo do tipo numérico relativo à velocidade máxima atingida pelo atleta em determinada sessão. O valor máximo observado foi de 36,46 km/h e o valor mínimo 9,04 km/h. Já a média global da equipa foi de $26,86 \pm 3,88$ km/h. Na ordem dos valores *missing* anteriormente registados, neste caso ocorreu uma percentagem de 45% com um valor total de 3055.

Player Load 3D, é um atributo utilizado pelo fabricante GPS para quantificar a totalidade das acelerações aplicadas em qualquer dos 3 planos: frontal, sagital e transversal. Portanto quanto mais rotações, acelerações, desacelerações e saltos maior será o valor de *Player Load 3D* registado. São também valores contínuos do tipo numérico com uma média de equipa de $281,80 \pm 116,01$ g com um valor máximo de 783,86 g e valor mínimo de 1,31 g. A totalidade de valores *missing* reportada foi de 3043 aproximadamente 44% da amostra.

Total Accelerations, neste atributo encontramos valores contínuos do tipo numérico relacionados com o número total de acelerações superiores a 2 m/s^2 desempenhadas pelo atleta em determinada sessão. A média de valores registados foi $83,51 \pm 40,13$ com valor máximo registado de 219 e valor mínimo 1. Já os valores *missing* registados foram de 2820 perfazendo 41% da totalidade da amostra.

Total Decelerations, ao contrário do parágrafo anterior, este atributo está relacionado com o número total de ações de desaceleração superiores a 2 m/s^2 desempenhadas pelo atleta. É também ele um atributo contínuo do tipo numérico com valor médio de $73,81 \pm 38,11$, máximo de 226 e mínimo 1. Relativamente aos valores *missing* foi registado mais um caso do que no seu atributo oposto.

Max Sprints, é, portanto, o atributo contínuo do tipo numérico indicador do número total de *sprints* com velocidade superior a 25km/h. A média da equipa é de $5,80 \pm 5,23$ com valor máximo de 46 e mínimo de 1. A contagem total de valores *missing* foi de 3980 e como justificado anteriormente esta alta

percentagem (58%) poderá estar associada à incapacidade dos atletas de atingirem velocidades nesta ordem de grandeza frequentemente.

Tabela 2: Tabela *load* pré processamento

Data	Semanas	Jogador	Posição	Peso (kg)	Poso (A)	Estado	RPE	totaldistance (0-1)	totaldistance (20-25)	totaldistance (7-14)	totaldistance (14-20)	Max Speed	PlayerLoad (N)	totaldistance (1-25)	total_acceleration	total_acceleration
18/08/2018	Semana 1 A01	GR	91	Campo	5											
18/08/2018	Semana 1 A02	GR	91	Campo	6											
18/08/2018	Semana 1 A03	GR	90	Campo	5											
18/08/2018	Semana 1 A04	DL	66	Campo	6	4365,020242	2973,742123	136,764693	865,1847240	384,1575894			17	7	101	87
18/08/2018	Semana 1 A05	DL	74	Campo	7	1278,208268	1163,989969		113,7802484	0,437610853						2
18/08/2018	Semana 1 A06	DC	83	Campo	6	2050,662320	740,1187796		210,5404238							
18/08/2018	Semana 1 A07	DC	82	Campo	5,5	1153,946186	1043,885647		70,06579807							
18/08/2018	Semana 1 A08	DC	80	Campo	6	5943,31195	5143,403841	104,1553000	180,156455	534,1905128						94
18/08/2018	Semana 1 A09	DC	80	Campo	6	5111,880293	2862,777096	134,3487009	158,511196	618,5237804						75
18/08/2018	Semana 1 A10	DL	79	Campo	7,5											
18/08/2018	Semana 1 A11	DL	85	Campo	6	4900,30913	3543,271816	104,0138258	1037,313191	395,3171485						74
18/08/2018	Semana 1 A12	MC	80	Campo	6,5	1035,18208	925,91824		109,257156							
18/08/2018	Semana 1 A13	MC	78	Campo	6	374,1184618	404,919875		344,960485	0,21377887						
18/08/2018	Semana 1 A17	MC	89	Campo	7	3729,355456	3373,134939	134,6395467	1553,175785	658,9113003						76
18/08/2018	Semana 1 A18	MC	81	Laboratório (NOC)												
18/08/2018	Semana 1 A19	MC	74	Campo	6,5	6254,516818	2388,371616	178,7425211	2894,919979	788,4447796						90
18/08/2018	Semana 1 A20	EX	83	Campo	7,5											
18/08/2018	Semana 1 A21	EX	72	Campo	7	709,7446843	181,2798824		124,4884019							
18/08/2018	Semana 1 A22	EX	68	Campo	7	5151,763043	2891,981038	181,1237713	1478,506188	548,957325						88
18/08/2018	Semana 1 A23	EX	77	Campo	7											
18/08/2018	Semana 1 A24	AV	83	Campo	7	971,4748788	861,7994879		89,67539089							
18/08/2018	Semana 1 A25	AV	80	Campo	7	5048,121132	3211,168144	124,1136051	1921,669308	943,6925768						81
18/08/2018	Semana 1 A26	GR	91	Campo	5											
18/08/2018	Semana 1 A27	GR	91	Campo	6											
18/08/2018	Semana 1 A28	GR	90	Campo	5											
18/08/2018	Semana 1 A29	DL	66	Campo	6	3960,213351	2957,241343	7,52733966	856,9874045	158,4512639						89
18/08/2018	Semana 1 A30	DL	74	Campo	7	3491,888131	3187,531244	8,567399207	2039,394707	480,6948046						32
18/08/2018	Semana 1 A31	DC	83	Campo	6	3684,446819	3288,517482	27,01010201	1913,984742	578,8770173						72
18/08/2018	Semana 1 A32	DC	82	Campo	5,5	6092,521577	3129,832082	16,02578994	2288,186196	658,4775986						48
18/08/2018	Semana 1 A33	DC	80	Campo	6	6676,712489	3293,139363	25,74991519	2612,082113	721,4997599						80
18/08/2018	Semana 1 A34	DC	83	Campo	6	6176,23047	3177,282166	64,4677596	2356,450613	579,9728348						83
18/08/2018	Semana 1 A35	DL	79	Campo	7,5	6218,067225	3521,432294	98,7605177	1990,117419	797,0204079						100
18/08/2018	Semana 1 A36	DL	85	Campo	6											
18/08/2018	Semana 1 A37	MC	80	Campo	6,5	6833,224279	3074,38376	23,47813654	2787,801441	947,5891632						88
18/08/2018	Semana 1 A38	MC	78	Campo	6	6277,295916	3463,351296	15,22976554	2142,264805	254,3708919						86
18/08/2018	Semana 1 A39	MC	89	Campo	7	6665,1455	2884,679486	43,37424337	2767,649501	873,442688						85
18/08/2018	Semana 1 A40	MC	81	Laboratório (NOC)												
18/08/2018	Semana 1 A41	MC	74	Campo	6,5	7390,152768	2934,070825	36,42800041	3354,451284	1084,350881						89
18/08/2018	Semana 1 A42	EX	83	Campo	7,5	6795,300487	2896,506647	35,73735556	3146,639741	749,7964934						82
18/08/2018	Semana 1 A43	EX	72	Campo	7	6037,688056	3371,519472	46,79446613	1603,879427	596,565518						77
18/08/2018	Semana 1 A44	EX	68	Campo	7	3648,100954	3301,400007	40,74351924	1746,319375	563,4215968						50
18/08/2018	Semana 1 A45	EX	77	Campo	7	6152,760094	3190,088863	68,88887273	2332,16245	660,3908921						100
18/08/2018	Semana 1 A46	AV	80	Campo	7											

Concluída a análise da tabela *load* repito o mesmo procedimento para a tabela *force*. Esta tabela é constituída por 13 diferentes atributos e 1308 entradas.

Como verificado na descrição anterior, o atributo Data é também nesta tabela o primeiro. O primeiro valor observado é de 27 de agosto de 2018 e o último 8 de maio de 2019, sendo que a duração total do ensaio foi de 254 dias. A discrepância nas datas de início dos registos verificada nas duas tabelas está relacionada com a disponibilidade dos equipamentos, apenas a 27 de agosto estiveram disponíveis os dispositivos para avaliação da força muscular exercida.

Jogador, também já descrito previamente neste capítulo, atua como chave estrangeira na ligação das duas tabelas.

Min Jogo, é um atributo contínuo do tipo numérico, onde, tal como o nome indica, são registados os minutos jogados por dado atleta. Como esperado o valor mínimo registado foi de 0 min e o valor máximo 97 min, a média da equipa foi de $46,70 \pm 37,63$ min. Foram registados 770 valores *missing* perfazendo um total de 58% da totalidade da amostra.

Jornada, regista a jornada associada a determinada semana de trabalho. É um atributo categórico do tipo texto de especial relevo para os profissionais da equipa técnica de modo a associarem a semana de trabalho a um determinado jogo.

Tempo, é também ele um atributo categórico do tipo texto no qual se regista o tempo em que determinada avaliação foi conduzida comparativamente

com a data do jogo. Neste atributo encontram-se dois valores distintos, MD+1 (*match day* + 1) com uma frequência de 356 casos e MD-2 (*match day* -2) com frequência de 952.

Squeeze 45°, é o primeiro atributo da tabela derivado dos valores captados pelos dispositivos de avaliação de produção de força muscular. Contem valores contínuos do tipo numérico onde a CVM é registada. Com o suporte do equipamento Smart Groin Trainer, o atleta realiza uma contração dos adutores com as pernas fletidas a 45°. O valor máximo registado foi de 948,28 N e o valor mínimo 143,17 N sendo a média do plantel $546,67 \pm 112,18$ N. Relativamente aos valores *missing*, estes perfazem 55% da totalidade da amostra com uma contagem total de 727. Como argumentado anteriormente, a volatilidade do trabalho com equipas profissionais de futebol, ou outros desportos de equipa, tornam o cumprimento de protocolos extremamente difícil. A escassez de equipamentos e também recursos humanos tornam o processo moroso, como tal, nem sempre o cumprimento das avaliações protocoladas foi possível. Como verificado nas grandes diferenças de frequência de leituras entre MD+1 e MD-2, é possível verificar que no dia após o jogo a grande prioridade da equipa técnica é a recuperação física dos atletas, por consequência, muitas das avaliações físicas passaram para segundo plano.

Ísquio Esquerdo, atributo contínuo do tipo numérico. São valores obtidos pelo teste isométrico de extensão do joelho esquerdo a 30° no equipamento Acceleration Leg Curl/Extension. O valor máximo observado foi de 703,14 N e o mínimo 256,76 N com uma média de $452,12 \pm 83,19$ N. Tal como no protocolo de *Squeeze 0°* foi registada uma grande percentagem de valores *missing*, aproximadamente 90% com uma contagem de 1180.

Ísquio Direito, descreve o mesmo processo relatado no parágrafo anterior, mas neste caso para o joelho direito. Os valores *missing* observados foram os mesmos, mas a média subiu para $466,79 \pm 89,78$ N, valor máximo 722,75 N e valor mínimo 279,49.

Ísquio Total, não é mais do que a soma dos valores recolhidos em cada um dos testes. Como tal, mais uma vez os valores *missing* mantêm-se, a média registada foi de $897,77 \pm 184,74$ N, valor máximo 1395,48 N e valor mínimo 339,31 N.

Nórdicos Esquerdo, é um atributo contínuo do tipo numérico recolhido com o equipamento Nordic Hamstring Addapter em exercício dinâmico de extensão dos isquiotibiais. Relativo à perna esquerda o valor máximo registado foi de 686,46 N e o mínimo de 205,94 N. Já a média foi de $442,19 \pm 81,63$ N. A

4. Metodologia

No âmbito desta dissertação pretende-se avaliar a eficácia e geral praticidade da monitorização dos níveis de produção de força em grupos musculares de referência como ferramenta de monitorização de fadiga em futebolistas de alta competição. Deste modo, o clube e *staff* ficaria dotado de uma ferramenta capaz de assistir à gestão do plantel, dado que, a fadiga é um fator major na produção atlética e desportiva como descrito no capítulo de estado da arte.

Durante a época desportiva foram recolhidos e analisados continuamente dados relativos ao posicionamento no campo obtidos através do sensor de GPS e dados de produção de força obtidos pelos equipamentos Smart Groin Trainer e Acceleration Leg Curl/Extension. Após esta recolha foi construído um *dataset* anteriormente descrito.

Utilizando metodologias de caracterização abordo os dados de força recolhidos 48h antes do jogo (MD-2) e 24h após jogo (MD+1) onde procuro diferenças entre os níveis de força exercidos nos dois momentos.

A dissertação tem os seguintes objetivos específicos:

- Descrição detalhada dos processos de recolha de dados constituintes do *dataset*;
- Descrição dos métodos de pré-processamento e processamento aplicados;
- Descrição detalhada do *dataset* considerado para o âmbito desta dissertação;
- Descrição dos métodos de caracterização de variáveis utilizados;

4.1 Metodologia de aquisição de dados

No decorrer da época desportiva de 2018/2019 foram adquiridos dados relativos a dois grandes grupos de trabalho distintos. Através da utilização do sensor de GPS cada atleta foi monitorizado nas suas movimentações dentro de campo, com os equipamentos Smart Groin Trainer e Acceleration Leg Curl/Extension foram recolhidos os valores de força exercida nos grupos musculares adutores, quadríceps e isquiotibiais.

Começando pelos dados relativos às movimentações dos futebolistas estes seguiram a seguinte metodologia de recolha para a construção do dataset pré processado, durante a totalidade da época. O movimento dos 30 atletas, com média de idade $25,74 \pm 4,89$ e média de pesos $79,65 \pm 7,86$ kg foi capturado utilizando 30 sensores de posicionamento tridimensionais Johan-Sports, um para cada jogador, com aquisição de dados a 5Hz,^[64] acoplados a um pequeno colete, os atletas vestem o equipamento e executam as rotinas de treino ou jogo pré-programadas.^[65]



Figura 15: Atleta do FCF utilizando o sensor de GPS Johan Sports

Em todos os períodos de recolha de dados a antena do equipamento foi situada junto à linha lateral em posição central de modo a uma receção de sinal dos satélites clara. Através deste equipamento foram obtidos algumas métricas disponíveis na aplicação Johan-Sports Analytics^[66], dentro as quais, *Total Distance*, *Walk Distance*, *Sprint Distance*, *Jog Distance*, *Run Distance*, *Max Speed*, *Player*

Load 3D, High Sprint Distance, High Sprints, Total Accelerations e Total Decelerations, já descritas no capítulo de descrição do *dataset* pré-processado.

Relativamente aos dados de aplicação de força, a metodologia de recolha destes dados foi dividida pelos três diferentes grupos musculares avaliados. No entanto a periodicidade de avaliação protocolada foi a mesma, sendo que todas as intervenções foram realizadas antes dos momentos de treino 48h antes do jogo e 24h após a sua conclusão.

Antes da avaliação os participantes realizaram um aquecimento ligeiro de 5 minutos de corrida de intensidade baixa e 8 minutos de alongamento dos membros inferiores e tronco [67]. Uma semana antes do início das avaliações, todos os atletas foram familiarizados com o tipo de exercícios a que iriam ser submetidos, de modo a diminuir possíveis erros de aprendizagem. Todos os testes foram efetuados por profissionais com anos de experiência na área médica e de performance desportiva. Para os testes de CVM dos adutores da bacia, cada participante posicionou-se em decúbito dorsal com a anca posicionada a 45° flexão e as pernas fletidas em 90°. [68] O valor máximo da contração foi quantificado pelo dinamómetro portátil Smart Groin Trainer, Neuroexcellence. A fiabilidade das medições realizadas por este equipamento pode ser encontrada em (Mesquita et al., 2018) [69]. O equipamento foi então colocado no ponto mais proeminente dos côndilos mediais do fémur. Os atletas foram instruídos a apertar o dispositivo com a máxima força durante 5 s.



Figura 16: Avaliação *Squeeze test* 45°

Para avaliação da força exercida pelos grupos musculares isquiotibiais os mesmos atletas foram avaliados uma perna de cada vez com o equipamento Acceleration Leg Curl/Extension. Os futebolistas foram instruídos a exercer força máxima de extensão do joelho em posição sentada durante um período de 5s.^[70]



Figura 17: Avaliação isométrica de extensão do joelho a 30°

Por fim, para avaliação dos grupos musculares isquiotibiais foi utilizado o acessório Nordic Hamstring Adapter for Acceleration Leg Curl/Extension. Os atletas foram instruídos a posicionarem-se de joelhos no topo do equipamento, fixando os seus tornozelos às células de carga monoaxiais aí posicionadas. Após indicação do profissional os futebolistas projetam-se lentamente para um plano anterior até à extensão máxima permitida pelo equipamento.^[71]



Figura 18: Avaliação de exercício de isquiotibiais Nórdico

4.2 Métodos de processamento do *dataset*

Neste capítulo descrevo as metodologias e racional aplicado no processamento do *dataset* consolidado pelas tabelas *force* e *load*.

Utilizando a ferramenta *Microsoft Excel 2016* e o *software Rapid Miner v.8.2.0* obtive a descrição apresentada no capítulo 3.

Tabela 4: Tabela *force* pós processamento

Row No.	Data	Jogador	Min Jogo	RPE	Jornada	Estado do C...	Tempo	Squeeze 45°
1	Aug 29, 2018 ...	A06	90	10	(4) Arouca C	Bom	MD-2	436.396
2	Aug 29, 2018 ...	A21	18	8	(4) Arouca C	Bom	MD-2	454.048
3	Sep 11, 2018...	A05	65	7	AMIG (VSC)	Bom	MD-2	583.496
4	Sep 11, 2018...	A06	0	0	AMIG (VSC)	Bom	MD-2	254.973
5	Sep 11, 2018...	A07	65	8	AMIG (VSC)	Bom	MD-2	554.076
6	Sep 11, 2018...	A08	25	8	AMIG (VSC)	Bom	MD-2	428.551
7	Sep 11, 2018...	A09	90	8	AMIG (VSC)	Bom	MD-2	668.814
8	Sep 11, 2018...	A10	65	7	AMIG (VSC)	Bom	MD-2	495.236
9	Sep 11, 2018...	A11	25	8	AMIG (VSC)	Bom	MD-2	542.308
10	Sep 11, 2018...	A12	0	6	AMIG (VSC)	Bom	MD-2	385.401
11	Sep 11, 2018...	A13	34	8	AMIG (VSC)	Bom	MD-2	428.551
12	Sep 11, 2018...	A15	56	7	AMIG (VSC)	Bom	MD-2	461.893
13	Sep 11, 2018...	A16	65	8.500	AMIG (VSC)	Bom	MD-2	545.250
14	Sep 11, 2018...	A17	65	8	AMIG (VSC)	Bom	MD-2	648.220
15	Sep 11, 2018...	A19	25	6.500	AMIG (VSC)	Bom	MD-2	469.739
16	Sep 11, 2018...	A20	90	9	AMIG (VSC)	Bom	MD-2	538.385
17	Sep 11, 2018...	A21	34	6	AMIG (VSC)	Bom	MD-2	511.907
18	Sep 11, 2018...	A14	65	7	AMIG (VSC)	Bom	MD-2	456.990

Na tabela *load*, foram inseridos manualmente os valores relativos à força exercida pelos atletas em diferentes grupos musculares, sob as diferentes nomenclaturas de atributos anteriormente apresentadas. Primeiramente, após a estruturação da tabela, os dados exportados pelos dispositivos foram convertidos da unidade Kg para a unidade de força do Sistema Internacional N.

No entanto, como argumentado anteriormente, nem todos os atributos se mostraram relevantes para o âmbito desta dissertação dado o extenso incumprimento dos protocolos de avaliação. Portanto na tabela *load* os atributos, Ísquio Esquerdo, Ísquio Direito, Ísquio Total, Nórdicos Esquerdo, Nórdicos Direito e Nórdicos Total foram removidos do *dataset* em consideração nesta dissertação.

Após este processamento, os valores *missing* foram também desconsiderados e eliminados, resultando assim numa tabela com 512 entradas e os atributos:

Data, Jogador, Minutos de Jogo, RPE, Jornada, Estado do Campo, Tempo e *Squeeze* 45°.

Foram recolhidas 332 medições de *Squeeze* 45° 48h antes de situações de jogo e 180 24h após o seu término.

O valor máximo de *Squeeze* 45° registado foi de 949,26 N e o valor mínimo 143,17 com um valor médio de $553,61 \pm 113,59$ N.

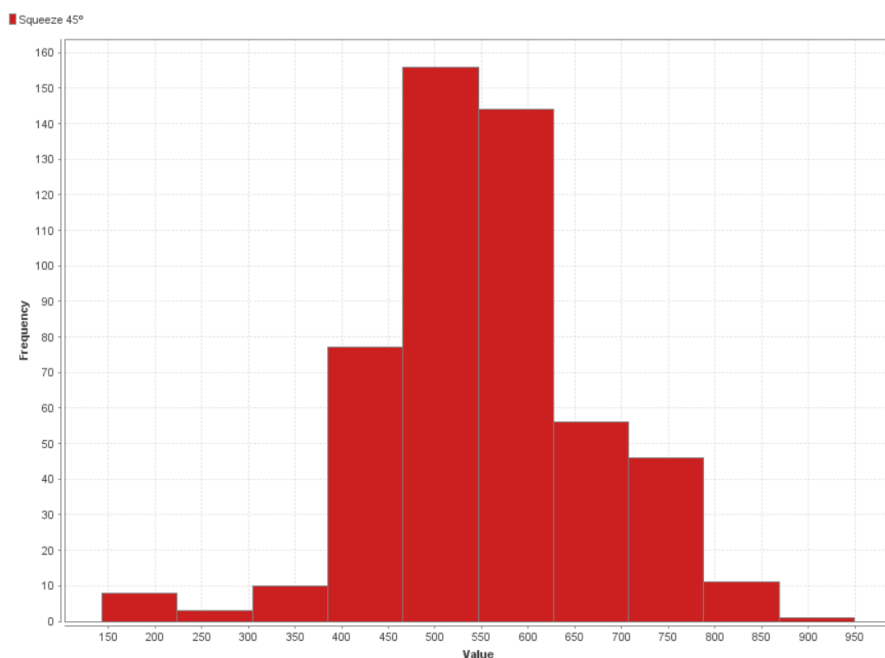


Figura 19: Distribuição dos valores de *Squeeze* 45°

Relativamente à tabela *load*, foram também removidos os valores *missing* causados pela inutilização dos sensores GPS, de uma amostra anterior de 6783 valores foram removidos 3262 valores perfazendo assim a totalidade de 3521 entradas no *dataset*.

Neste novo *dataset* temos então a seguinte análise estatística das métricas recolhidas pelo sensor GPS:

Tabela 5: Análise estatística da tabela *load*

Atributo	Média	Desvio Padrão	Valor Máximo	Valor Mínimo	Valores <i>Missing</i>
<i>Total Distance</i>	5584,86 m	2363,62 m	15767,78 m	20 m	0
<i>Walk Distance</i>	2857,99 m	874,29 m	6808,25 m	66,77 m	21
<i>Jog Distance</i>	1909,05 m	1131,64 m	7696,22 m	1,43 m	16
<i>Run Dustance</i>	642,30 m	539,66 m	3412,07 m	0,21 m	58
<i>Sprint Distance</i>	157,39 m	160,98 m	961,95 m	0,42 m	202
<i>High Sprint Distance</i>	53,92 m	63,71 m	381,52 m	0,07 m	1078
<i>Max Speed</i>	26,83 m	3,87 m	36,47 m	9,03 m	255
<i>Player Load 3D</i>	284,66 g	117,93 g	783,86 g	1,31 g	244
<i>High Sprints</i>	5,81	5,18	46	1	1125
<i>Total Decelerations</i>	88,72	40,05	219	1	50
<i>Total Accelerations</i>	74,60	38,39	226	1	52

4.3 Descrição da população de estudo

O plantel do FCF na época desportiva 2018/2019 foi constituído por uma totalidade de 30 atletas ($n=30$), no entanto, nem todos eles representaram o clube durante a total duração do estudo ou cumpriram os requisitos de inclusão.

Aplicando os seguintes critérios de inclusão:

- Atletas do plantel sénior FCF;
- Atletas pertencem aos quadros do clube durante a totalidade da época;
- Atletas sem lesões de longa duração ($t > 2$ meses)
- Atletas não Guarda Redes;

Vemos assim reduzido o número de participantes para 20 atletas. Os futebolistas (A01, A02, A03) são excluídos por serem os guarda-redes do clube, (A22, A26 e A04) foram dispensados no mercado de janeiro e (A27, A28, A29 e A30) foram apenas integrados no plantel na abertura em meados da época.

A média de idades registada da população de estudo foi de $25,7 \pm 4,64$ com média de pesos de $78,8 \pm 6,55$ kg.

A distribuição dos atletas pelas posições que ocupam no terreno foi a seguinte:

- DC, defesas centrais, 4 atletas
- DL, defesas laterais, 4 atletas
- MC, médios centro, 7 atletas
- EX, extremos laterais, 3 atletas
- AV, avançados, 2 atletas

4.4 Metodologia de análise

Com o objetivo de entender a capacidade de monitorização de fadiga muscular periférica através da monitorização da força exercida foi construído um relatório dinâmico de análise de dados com a utilização do *software* da Microsoft, Powerbi v.2.69.

Como descrito anteriormente, por incumprimento nas metodologias de recolhas de dados os atributos relativos à produção de força nos grupos musculares isquiotibiais e quadríceps femorais foram removidos, assim, a análise irá focar-se na produção de força máxima nos adutores em exercício de *Squeeze* 45°.

O contributo dos adutores no contexto do futebol é de vital importância dado o seu papel fundamental em ações de remate, sprint e mudanças de direção. É também este grupo muscular um dos mais fustigados por lesões nos futebolistas, assim sendo, a monitorização próxima dos níveis de produção de força pode estar relacionada com uma melhor preparação física do atleta bem como um menor risco de lesão.^{[72][73][74]}

Como tal, é expectável que através da monitorização da força máxima registada 48h antes do jogo e 24h após se detetem diminuições na globalidade do plantel.

O trabalho isométrico desenvolvido nestas ações de avaliação é também um grande contribuidor no condicionamento das estruturas muscular^[68], como tal é também analisado a evolução dos níveis de força em ambos os momentos de avaliação durante a época desportiva.

Em simultâneo, são também analisados os dados obtidos pelo sensor de GPS de modo a inferir sobre o nível de capacidade física produzida pela equipa quando comparando com equipas de elite Europeu e de que modo os atletas experienciam os estímulos de treino quando comparados com os estímulos de jogo. Deste modo será possível inferir sobre a capacidade de mimetização dos estímulos de jogo em momento de treino e da mesma forma relacionar estes estímulos mais fatigantes com o nível de preparação dos atletas.

No *software* Powerbi, foi criada a seguinte base de dados relacional representada na imagem:

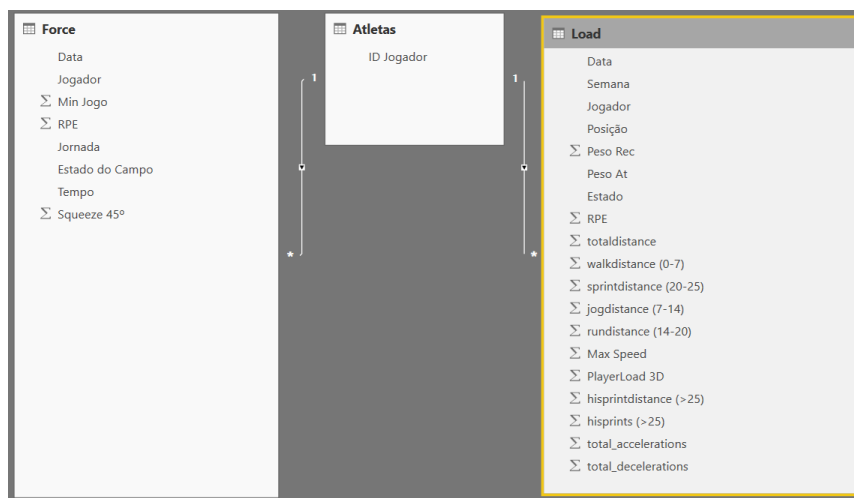


Figura 20: Modelo relacional do *dataset*

Através de metodologia de análise de caracterização foram avaliadas as relações entre a força da CMV 48h antes do jogo e 24h após a evolução da força exercida em CMV em ambos os casos durante a época desportiva de 2018/2019, a intensidade dos treinos em comparação com momentos de jogo através da análise do número de acelerações e desacelerações, a distância total percorrida e a distância de *High Sprint* em ambos os casos.

5. Resultados

Analisando os dados obtidos pela CMV no teste *Squeeze* 45° não foram observadas diferenças estatisticamente significativas nas medições efetuadas no período de 48h antes do jogo quando comparadas com as medições efetuadas 24h após. A média observada nas medições pré jogo foi de $562,83 \pm 117,45$ N e no período pós jogo $551,00 \pm 105,44$ N. A mediana registada foi 562,83 N e 548,19 N, respetivamente. Já os primeiros e terceiros quartis registaram valores de 467,85 N e 649,94 N, 473,66 N e 621,25 N, respetivamente. A sua representação gráfica poderá ser observada na seguinte imagem.

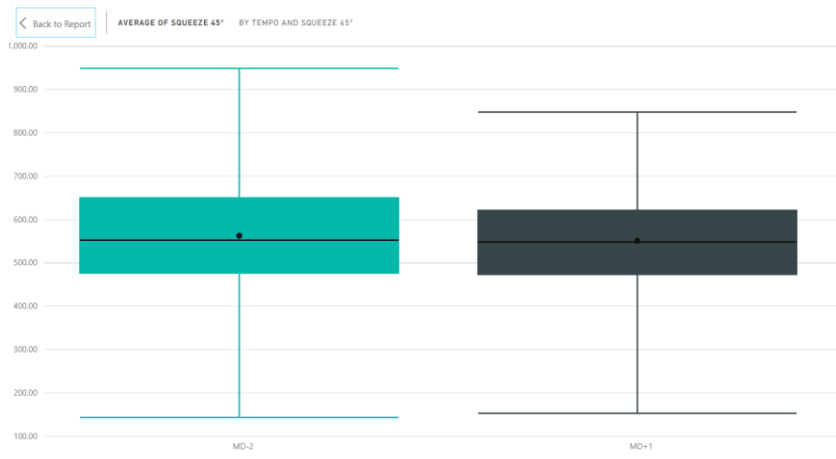


Figura 21: Boxplot de caracterização de força exercida em *Squeeze* 45° em MD-2 e MD+1

Relativamente à evolução temporal da CMV nos adutores foi observado um aumento na produção de força em ambos os casos de estudo. No início da época a média da CMV registada era de $445,22 \pm 8,83$ N no MD-2, já no caso do MD+1 a média observada foi de $502,59 \pm 97,91$ N. As mesmas avaliações (MD-2 e MD+1) registaram uma média de $569,77 \pm 88,61$ N e $559,31 \pm 72,69$ N, respetivamente, perfazendo assim uma evolução de 21,85% e 10,14% na produção de força pelos grupos musculares adutores.

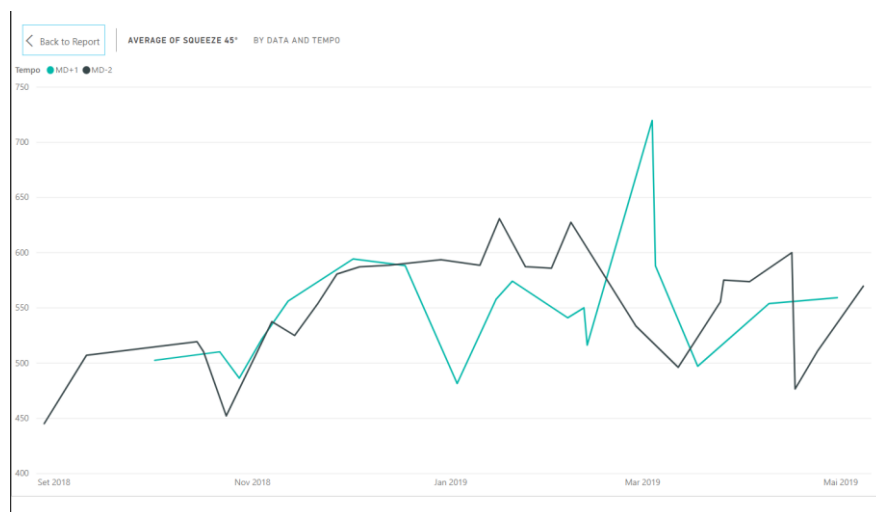


Figura 22: Gráfico linear de progressão dos níveis de força em *Squeeze* 45° durante a época 2018/2019

Observando agora os resultados obtidos na análise das métricas derivadas do sensor de GPS é possível identificar que os estímulos de jogo são em média mais intensos do que aqueles registados em contexto de treino. Uma análise extensiva aos parâmetros de alta intensidade (ações de aceleração e desaceleração, distância percorrida em *sprint* máximo e distância total percorrida), fatores estes associados à fadiga muscular nos adutores, mostra que para o caso de ações de aceleração em situação de jogo a média da contagem foi de $104,53 \pm 47,85$ e em situações de treino $90,87 \pm 36,07$. Registrando-se assim um aumento de 13,06% destas ações em situação de jogo.

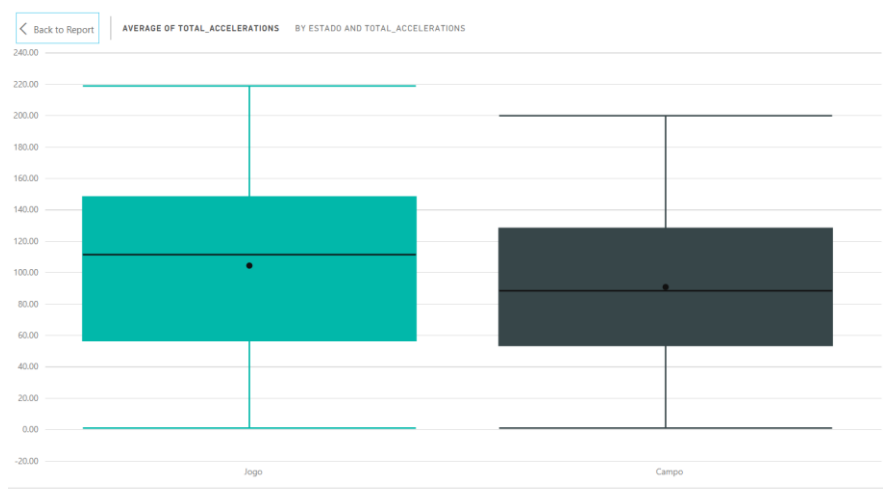


Figura 23: Boxplot de caraterização das ações de aceleração em jogo e treino

A distância em *High Sprint* apresenta também uma discrepância na ordem dos 65,62% entre as médias observadas em situações de jogo e de treino. A média registada em situações de jogo foi de $112,45 \pm 6,14$ m e em situações de treino $38,66 \pm 4,06$ m. O seguinte *boxplot* ilustra graficamente estas afirmações:

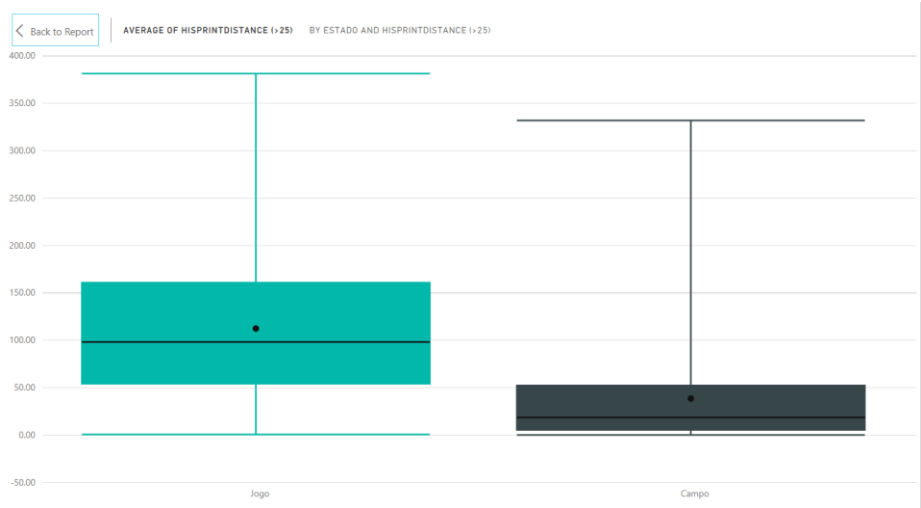


Figura 24: Boxplot de caraterização das ações de *High Sprint* em jogo e treino

Por último, foi também avaliada a distância total percorrida pelos atletas em ambos os momentos. Como esperado, a distância total percorrida em situações de jogo é vastamente superior à registada em situações de treino. Com uma média de $8,38 \pm 3,37$ km a distância percorrida pelos atletas em jogo foi 1,6 vezes maior do que em treino, $5,23 \pm 1,81$ km.

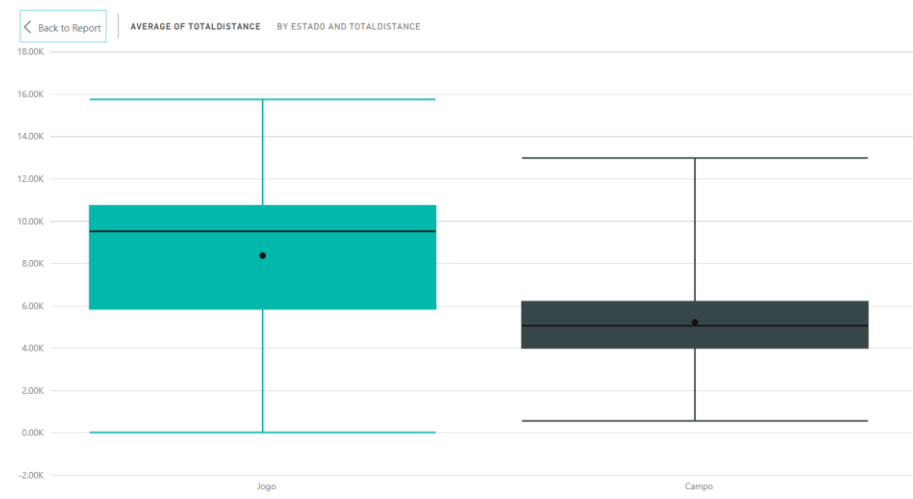


Figura 25: Boxplot de caracterização da distância total percorrida em situações de jogo e treino

6. Discussão

O objetivo desta dissertação foi analisar dados biométricos tais como os níveis de produção de força em grupos musculares chave e as movimentações dos atletas em contexto de jogo e treino para a quantificação de fadiga periférica e assim promover a utilização deste tipo de ferramentas para auxílio à equipa técnica na gestão do plantel. No entanto, não foi possível durante o decorrer da época desportiva de 2018/2019 recolher os dados de uma forma sistemática, tal como descrita nas metodologias de recolha dos dados. Vários fatores como a gestão do escasso tempo que a equipa de *Performance* e alto rendimento dispõe para trabalho de avaliação e ginásio e alterações na própria estrutura do clube foram determinantes no insucesso desta etapa.

No entanto, com maior critério, os dados do exercício de *Squeeze* 45° foram sistematicamente recolhidos de modo a terem relevo nesta dissertação.

Ao contrário do esperado, não existe aparente relação entre a força produzida pelos adutores nas avaliações realizadas antes e após o jogo como publicado por Row, Gregory A.B, et al 2016.^[72]

Uma aparente justificação para este acontecimento é a excelente preparação física e gestão de cargas dos atletas no decorrer da época, o performance atlético dos atletas está sempre dependente de um bom programa de treino sistematizado, sendo que, torna-se também fundamental considerar que a fadiga e os efeitos de treino são efeitos dinâmicos e não estáticos. Como tal, a monitorização próxima dos atletas poderá ter um efeito preventivo nos fenómenos de *overtraining* uma vez que, se os efeitos de treino ainda não se tiverem dispersado e o atleta estiver recuperado, voltar a treinar irá exacerbar esses efeitos fatigando ainda mais o atleta.^[73] A possibilidade de monitorizar e quantificar o esforço exercido pelos jogadores permitiu à equipa técnica ajustar e/ou adaptar as condições de treino individualmente para cada futebolista.

Como observado no capítulo anterior a média das ações de alta intensidade em nenhum caso foi maior em momentos de treino do que de jogo. A

prevenção de situações de treino excessivo é fundamental no desenvolvimento da capacidade física e capacidade geral do atleta executar as ações de jogo ao seu melhor nível nos momentos que realmente importam, os jogos.

Esta discrepância pode também ela ser justificada pela priorização de rotinas de treino em situações de campo reduzido, desse modo, todas as ações relativas às movimentações dos atletas são menores do que as registadas em situação de jogo.

Comparando os dados recolhidos em situações de jogo com as equipas de elite Europeu, é possível observar que as exigências da Segunda Liga Portuguesa são comparativamente mais baixas do que as registadas em jogos entre equipas de topo. Observando o estudo publicado Marcin Andrzejewski, et al. 2015,^[62] concluímos que a média de distância total percorrida pelos jogadores do FCF ($8,38 \pm 3,37$ km) é 1,32 vezes inferior à distância percorrida pelos atletas de equipas de elite Europeu ($11,06 \pm 0,79$ km), assim como a distância de *High Sprint* que no caso dos atletas do FCF foi 2,10 vezes inferior à registada pelos atletas em comparação ($112,45 \pm 6,14$ e 237 ± 123 m) respetivamente. Estes resultados poderão não ser indicadores de uma má preparação física, podem estar simplesmente relacionados com a capacidade da equipa se adaptar aos estímulos de jogo e durante esse período intuitivamente gerirem o seu esforço.

Como descrito extensamente no capítulo 2, os métodos para quantificação de fadiga são ainda ao dia de hoje extremamente complexos e difíceis de validar. A própria definição de fadiga e a sua monitorização ótima não é coerente na comunidade científica e profissional, como tal, apesar dos resultados inconclusivos apresentados, considero que a possibilidade da equipa técnica de monitorizar e acompanhar diariamente com dados concretos os níveis de força dos seus atletas é uma mais valia como uma ferramenta de gestão, tanto para os profissionais associados ao gabinete de alto rendimento como a para a equipa técnica que poderá tomar decisões mais ponderadas na atribuição de minutos a determinado jogador.

6.1 Trabalho Futuro

Finalizada a época desportiva de 2018/2019 termina também o projeto inserido no âmbito desta dissertação, no entanto, aguardo novo convite para o melhoramento e desenvolvimento de novas técnicas a serem aplicadas na época de 2019/2020.

Espera-se nesta nova época integrar novos dados biométricos, como a variabilidade da frequência cardíaca em jogo e treino, questionários de bem-estar e avaliações de força aos músculos da canela, solear e gastrocnémios com um aperfeiçoamento das práticas utilizadas este ano.

Neste contexto, pretende-se também desenvolver uma plataforma de gestão Web onde todos os profissionais de interesse dentro do clube poderão consultar e acrescentar informação relativa aos atletas..

Concluo esta dissertação com um profundo sentimento de satisfação pela possibilidade de colaborar com os excelentes profissionais do FCF numa área que me é tão especial e com a esperança de no futuro poder voltar a fazer parte deste projeto.

Referências

- [1] Dias A. Incidência e padrões de lesão em futebol profissional, um estudo durante três épocas consecutivas com uma equipa de elite. Faculdade de Desporto da Universidade do Porto; 2011.
- [2] Hagglund M, Waldén M, Magnusson H, Kristenson K, Bengtsson H, Ekstrand J. Injuries affect team performance negatively in professional football: An 11-year old follow-up of UEFA Champions League injury study. *Br J Sports Med.* 2013.
- [3] Ekstrand J, Hagglund M, Waldén M. Injury incidence and injury patterns in professional football: The UEFA injury study. *Br J Sports Med.* 2011.
- [4] Real cost of injuries in professional sport [Internet]. Kitman Labs. Available from: <https://www.kitmanlabs.com/knowledge/real-cost-of-injuries-in-professional-sport&type=subject&tag=Analytics>
- [5] Taylor J., Amann M., Duchateau J., Meeusen R., Rice C.. Neural contributions to muscle fatigue: from the brain to the muscle and back again. *Medicine & Science in Sports & Exercise.* 2016.
- [6] Asmussen E., Muscle fatigue. *Medicine and Science in Sports.* 1979
- [7] .Enoka R. M., Stuart D. G., Neurobiology of muscle fatigue. *Journal of Applied Physiology.* 1992.
- [8] Bigland-Ritchie B., Rice C. L., Gargland S. J., Walsh M. L. (1995), Task-dependent factors in fatigue of Human voluntary contractions. In: Gandevia S.C., Enoka R.M., McComas A.J. Stuart D.G. Thomas C.K., Pierce P.A. (eds) *Fatigue. Advances in Experimental Medicine and Biology*, vol 384 Springer, Boston, Ma
- [9] Baudry, S., Klass, M., Pasquet, B. et al. *Eur J Appl Physiol* (2007) 100:515
- [10] Hunter S. K., Chritchlow A., Enoka R. M., Muscle endurance is greater for old men compared with strength-matched young men, *Journal of Applied Physiology*, 2005.

- [11] Hicks AL., Kent-Braun J. & Ditor D.S., Sex differences in human skeletal muscle fatigue. *Exerc Sport Sci Rev* 29, 2006.
- [12] Clark B.C., Collier S.R., Manini T.M., Ploutz-Snyder L.I., Sex differences in muscle fatigability and activation patterns of the human quadriceps femoris. *Eur J Appl physiol*, 2005.
- [13] de Ruiter C. J., Goudsmit J. F., Van Tricht J. A., de Haan A., The isometric torque at which knee-extensor muscle reoxygenation stops, *Med Sci Sports Exerc*, 2007.
- [14] Hunter S. K., Critchlow A., Shin I. S., Enoka R. M., Fatigability of the elbow flexor muscles for a sustained submaximal contraction in similar men and women matched for strength, *J Appl Physiol*, 2004.
- [15] Russ D. W., Lanza I. R., Rothman D., Kent-Braun J. A., Sex differences in glycolysis during brief, intense isometric contractions. *Muscle Nerve*, 2005.
- [16] Carroll T. J., Taylor J. L., Gandevia S. C., Recovery of central and peripheral neuromuscular fatigue after exercise, *J Appl Physiol*, 2017.
- [17] Wernbom M., Augustsson J., Roland T., The influence of frequency, intensity, volume and mode of strength training on whole muscle cross-sectional area in Humans, *Sports Medicine*, 2007.
- [18] Dawson M. J., Gadian D. G., Wilkie D. R., Muscular fatigue investigated by phosphorus nuclear magnetic resonance, *Nature*, 1978.
- [19] Takada S., Okita K., Suga T., Omokawa M., Kadoguchi T., Low-intensity exercise can increase muscle mass and strength proportionally to enhanced metabolic stress under ischemic conditions, *J Appl Physiol*, 2012.
- [20] Tesch P. A., Colliander E. B., Kaiser P., Muscle metabolism during intense, heavy-resistance exercise, *Eur J Appl Physiol*, 1986.
- [21] Blinder-Macleod S. A., Russ D., Effects of activation frequency and force on low-frequency fatigue in human skeletal muscle. *J Appl Physiol*, 1999.
- [22] Chin E., Allen D., Role of intracellular calcium and metabolites in low-frequency fatigue of mouse skeletal muscle. *J Appl Physiol*, 1997.
- [23] Reid M., Muscle fatigue: mechanisms and regulation in CKSLPeO Hanninen, *Handbook of oxidants and antioxidants in exercise*, Amsterdam, Elsevier Science B. V., 2000.
- [24] Stienen G. Papp Z., Zaremba R., Influence of inorganic phosphate and pH on sarcoplasmic reticular ATPase in skinned muscle fibers of *Xenopus laevis*, *J Physiol*, 1999.

- [25] Essing D., Nosek T., Muscle fatigue and induction of stress protein genes. A dual function of reactive oxygen species?, *J App Physiol*, 1997.
- [26] Segersted O., Sjogaard G., Dynamics and consequences of potassium shifts in skeletal muscle and heart during exercise. *Physiological Reviews*, 2000.
- [27] Bangsbo J. Madsen K. Kiens B., Richter E., Effect of muscle acidity on muscle metabolism and fatigue during intense exercise in man, *J Physiol*, 1996.
- [28] Bangsbo J., Physiology of muscle fatigue during intense exercise, *The clinical pharmacology of sport and exercise*, Elsevier Science, 1997.
- [29] Sundstrup E., Jakobsen M. D., Andersen C. H., Zebis M. K., Mortensen O. S., Andersen L. L., Muscle activation strategies during strength training with heavy loading vs. repetitions to failure, *J Strength Cond Res*, 2012.
- [30] Jakobsen M. D., Sundstrup E., Andersen C. H., Zebis M. K., Mortensen P., Andersen L.L., Evaluation of muscle activity during a standardized shoulder resistance training bout in novice individuals, *J Strength Cond Res*, 2012.
- [31] Stackhouse S., Dean J., Lee S., Blinder-Macload S., Measurement of central activation failure of the quadriceps femoris in healthy adults. *Muscle and Nerve*, 2000.
- [32] Davis M., Fitts R., Mechanisms of muscular fatigue, *ACSM'S resource manual – guidelines for exercise testing and prescription*, Baltimore, Lippincott Williams & Wilkins, 2001.
- [33] Gandevia S., Spinal and supraspinal factors in human muscle fatigue. *Physiol Rev*, 2001.
- [34] Davis M, Central and peripheral factors in fatigue, *J Sports Sciences*, 1995.
- [35] Davis M., Bailey S., Possible mechanisms of central nervous system fatigue during exercise, *Med Sci Sports*, 1997.
- [36] Mohr M., Krustup P., Bangsboo J., Match performance of high-standard soccer players with special reference to the development of fatigue. *J Sports Sci*, 2003.
- [37] Saltin B., Metabolic fundamentals in exercise. *Med Sci Sports*, 2005.
- [38] Krstrup P., Mohr M., Steensberg A. et al, Muscle and blood metabolites during a soccer game: implications for sprint performance, *Med Sci Sports Exerc*, 2006.

- [39] Ekblom B., *Applied physiology of soccer*, Sports Med, 1986
- [40] Cheung S. S., Sleivert G.G., Multiple triggers for hyperthermic fatigue and exhaustion. *Exerc Sport Sci Rev*, 2004.
- [41] Mohr M., Krstrup P., Nybo L., et al., Muscle temperature and sprint performance during soccer matches: beneficial effect of re-warm-up at half time. *J Med Sci Sports*, 2004.
- [42] Rahnema N., Reilley T., Lees A., et al., Muscle fatigue induced by exercise simulating the work-rate of soccer. *J Sports Sci*, 2003.
- [43] Lees A., Davies T., The effects of fatigue on soccer kick biomechanics, *J Sports Sci*, 1988.
- [44] Nimmo M. A., Ekblom B., Fatigue and illness in athletes, *J Sports Sci*, 207.
- [45] Halson S. L., Monitoring training load to understand fatigue in athletes. *Sports Med*, 2014.
- [46] Virus A. M., Hackey A. C., Valja E., Karelson K., Janson T., Viru M., Influence of prolonged continuous exercise on hormone responses to subsequent exercise in humans, *J Appl Physiol*, 2001.
- [47] Gregson W., Drust B., Atkinson G., Salvo V. D., Match to match variability of high-speed activities in Premier League Soccer, *J Sports Med*, 2010.
- [48] Saw A. E., Main L. C., Gatin P. B., Monitoring the athlete training response: subjective self reported measures trump commonly used objective measures, a systematic review, *J Sports Med*, 2015.
- [49] Buchheit M., Racinais S., Bilsborough J. C., et al., Monitoring fitness, fatigue and running performance during a pre-season training camp in elite football players. *Sci Med Sport*, 2013.
- [50] Taylor K., Chapman D., Cronin J., Newton M., Gill N., Fatigue monitoring in high performance sport: a survey of current trends, *J Austr Strength*, 2012.
- [51] Andersson H., Raastad T., Nilsson J., Paulsen G., Garthe I., Kadi F., Neuromuscular fatigue and recovery in elite female soccer: effects of active recovery, *J Sport Med*, 2008.
- [52] Gibson N. E., Boyd A. J., Murray A. M., Countermovement jump is not affected during final competition preparation periods in elite rugby sevens players. *J Strength Cond Res*. 2016.

- [53] Gathercole R., Sporer B., Stellingwerff T., Sleivert G., Alternative countermovement jump analysis to quantify acute neuromuscular fatigue. *J Sports Physiol*, 2015.
- [54] Stanley J., Peake J. M., Buchheit M., Cardiac parasympathetic reactivation following exercise: implications for training prescription. *Sports Med*, 2013.
- [55] Thorpe R. T., Strudwick A. J., Buchheit M., Atkinson G. Drust B., Gregson W., Monitoring fatigue during the in-season competitive phase in elite soccer players, *J Sports Physiol Perform*, 2015.
- [56] Pews D., Laursen P., Kilding A., Buchheit M., Heart-rate variability and training intensity distribution in elite rowers. *J Physiol Perform*, 2014.
- [57] Ispirlidis I., Fatouros I. G., Jamurtas A. Z., et al., Time-course of changes in inflammatory and performance responses following a soccer game. *J Sport Med*, 2008.
- [58] Mohr M., Draganidis D., Chatzinikolaou, et al. Muscle damage, inflammatory, immune and performance responses to three football games in 1 week in competitive male players. *J Appl Physiol*, 2016.
- [59] Nédélec M., McCall A., Carling C., Legall F., Berthoin S., Dupont G., Recovery in soccer part I- post match fatigue and time course of recovery. *Sports Med*, 2012.
- [60] Han J., Kamber M., Pei J., 2000, *Data Mining, Concepts and Techniques*, Morgan Kaufman.
- [61] Atzeni P., *Relational database theory*, University of Milan.
- [62] Andrzejewski M., Chmura J., Pluta B., Konarski J., Sprinting activities and distance covered by top level Europa League soccer players, *I J Sports Science and Coaching*, 2015.
- [63] Moreno-Pérez V., Travassos B., Calado A., Skok O. G., Coso J., Villabuena A., Adductor squeeze test and groin injuries in elite football players: A prospective study, *Physical Therapy in Sport*, 2019.
- [64] Johan Sports Gps Hardware Specifications [Internet], Available at: <https://www.johan-sports.com/hardware/>
- [65] Jennings D., Cormack S., Coutts A. J., Boyd L., Aughey R. J., The validity and reliability of GPS units for measuring distance in team sport specific running patterns, *I J Sports Physiology and Performance*, 2010.
- [66] Johan Sports GPS Software Analytics [Internet] Available at: <https://app.johan-sports.com/#/faq>

- [67] Cejudo, A., Sainz de Baranda, P., Ayala, F., & Santonja, F., Test-retest reliability of seven common clinical tests for assessing lower extremity muscle flexibility in futsal and handball players. *Physical Therapy in Sport*, 2015.
- [68] Delahunt, E., Kennelly, C., McEntee, B. L., Coughlan, G. F., & Green, B. S., The thigh adductor squeeze test: 45 of hip flexion as the optimal test position for eliciting adductor muscle activity and maximum pressure values. *Manual therapy*, 2011.
- [69] Mesquita, R., Gonçalves, B., Tavares, F., Brito, J., Correia, P., Santos, P., & Mil-Homens, P, Maximal strength and rate of force development of hip adduction and abduction: reliability of measures from a portable dynamometer. Paper presented at the International Congress of Strength Training, Perth, Western Australia, 2018.
- [70] Deones V. L., Wiley S. C., Worrel T., Assessment of quadriceps muscle performance by a hand held dynamometer and an isokinetic dynamometer, *Journal of Orthopaedic & Sports Physical Therapy*, 1994.
- [71] Opar D. A., Piatkowski T., Williams M., Shield A. J., A novel device using the Nordic hamstring exercise to assess eccentric knee flexor strength: A reliability and retrospective injury study, *Journal of Orthopaedic & Sports Physical Therapy*, 2013.
- [72] Gregory A.B., Phibbs, Padraic J., et al., Changes in adductor strength after competition in academy rugby players, *J Strength and Conditioning Research*, 2016.
- [73] Robert H., et al., Lower extremity muscle activation and alignment during the soccer instep and side foot kicks, *Journal of Orthopaedic & Sports Physical Therapy*, 2007.
- [74] Moore S. P., Rushmer D. S., Windus S. L., et al., *Exp Brain Res*, 1988.

SEDE ADMINISTRATIVA

FACULDADE DE **MEDICINA**

FACULDADE DE **CIÊNCIAS**

