ORIGINAL ARTICLE

Relative Age Effects and the Premier League's Elite Player Performance Plan (EPPP): A Comparison of Birthdate Distributions Within and Between Age Groups

Greg Doncaster[1] · Adam L. Kelly[2] · Alexander B. T. McAuley[2] · Andrew Cain[3] · Mark Partington[1] · Lee Nelson[1] · Jimmy O'Gorman[1]

Abstract

The purpose of this study was to investigate the prevalence of relative age effects (RAEs) within and between U18, U21 and professional senior squads, that compete in the highest (respective) leagues within England. Birthdate, playing position and age (years) of U18 ($n = 487$), U21 ($n = 350$), and senior ($n = 396$) players from squads competing in the highest divisions of their respective age groups were obtained. Moreover, nationality (UK or Non-UK) was recorded for U21 and senior players, with estimated market value also obtained for senior players. Chi-square tests, Cramer's V and odds ratios with 95% confidence intervals were used to compare observed and expected birthdate distributions. A selection bias toward relatively older players was evident within U18 and U21 squads, across all positions. Furthermore, analysis of age bands within each age group also revealed an increase in the prevalence of RAEs throughout each age group. In contrast, analysis of senior squads revealed no significant deviations in birthdate distributions when considered as a whole sample, as separate age bands or by position. However, although non-significant, Q4 players were found to have the highest estimated market value. Results demonstrate RAEs are prevalent across U18 and U21 age groups at the highest level of competitive football within England, yet this is not representative of the birthdate distributions within senior squads. Ongoing research is needed to highlight the prevalence of RAEs within academies, particularly when these RAEs are not representative of birthdate distributions within professional senior squads.

Keywords Talent development · Talent identification · Team sports · Competition · Underdog hypothesis

Introduction

Football academies are responsible for identifying and developing young players to reach their potential, with the ultimate aim of achieving professional status to provide both sporting and financial success for their respective club [31]. In England, under the guise of the Elite Player Performance plan (EPPP, Premier [29]), professional football academies can sign individuals as young as aged 8 years—although many clubs are recruiting players much younger than this as part of their pre-academy programmes [15]. Early identification of talented youth football players can help maximize the possibility of providing (long-term) expert support to the *most* talented players. However, the complex nature of identifying those who have the talent to achieve expertise often results in a preference for current performance capabilities rather than long-term potential, which manifests in the prevalence of relative age effects (RAEs) (see [15] for an overview in football).

Relative age effects are a well-known phenomenon that refers to a distinct over-representation of players born earlier in the selection year (e.g., individuals born in September in England where the cut-off dates are September to August) for a given cohort [7]. The existence of a selection-bias toward players born earlier in the selection year is commonly reported within younger age groups across Europe [6, 18,

✉ Greg Doncaster
doncasg@edgehill.ac.uk

1 Department of Physical Activity and Sport, Faculty of Arts and Sciences, Edge Hill University, Ormskirk L39 4QP, UK

2 Athlete Development and Youth Sport (ADYS) Lab, Research Centre for Life and Sport Sciences (CLaSS), Birmingham City University, Birmingham, UK

3 Kitman Labs, Dublin, Ireland

27] and international (FIFA World Cups) football [26]. The ongoing impact of RAEs were highlighted by Helsen et al. [13], who observed no change in the prevalence (i.e., over-representation of relatively older players) over a 10-year period (from 2000–01 to 2010–11) across ten European countries. In addition, research has shown that RAEs vary across playing positions, with some studies displaying more pronounced RAEs in defensive and midfield positions [25, 32, 35] whilst others display an over-representation of relatively older players within attacking positions [28]. As such, research investigating RAEs should continue to give consideration toward playing position, allowing for an improved understanding and appreciation of the extent to which selection biases are prevalent across positions.

More specifically to England, RAEs were recognised in the development of the Premier Leagues EPPP in 2011 [29] as a *specific issue* (p. 59) and a proposed area of research within the future provision of the EPPP. Furthermore, coach education programmes (e.g. the FA Youth Award) have sought to include content regarding RAEs. Yet, since the introduction of the EPPP, developments in coach education programmes, and reformatting of the competitive leagues within the professional development phase (i.e., U18 North and South Premier Leagues and Premier League 2 competitions), there has been little insight into whether these have impacted upon the presence of RAEs, and to what extent any RAEs identified within youth age groups are indicative of birth-date distributions within elite senior teams. Therefore, the degree to which RAEs persist into older age groups (U18, U21 and senior squads) and the likelihood of being selected requires further investigation [17]. Indeed, while RAEs may already be present leading into U18 football, whether this remains both within (e.g., < 17 years old, 17–18 years old and > 18 years old, within the U18 squads) and across (U18, U21 and senior) age groups could help to inform future research and practice within football academies.

Previously, longitudinal research [17] examining the prevalence of RAEs across twelve seasons within the U18 squad of an English professional (Category 3, Tier 4) football academy reported that birthdate distributions were significantly skewed to those born earlier in the year (1st Quartile $n = 224$ [40.3%], 2nd Quartile $n = 168$ [30.2%], 3rd Quartile $n = 88$ [15.8%], 4th Quartile $n = 76$ [13.7%]). Although, Kelly et al. [17] also reported that of the 27 (7.4%) players who received a professional contract, there was a significantly greater proportion of those born later in the selection year relative to the pool of players available within each birth quartile (1st Quartile $n = 5$, 2nd Quartile $n = 8$, 3rd Quartile $n = 6$, 4th Quartile $n = 8$). As such, this has led to the suggestion of an 'underdog hypothesis', whereby being born later in the selection year (4th quartile) potentially facilitates an individual's long-term development, as these players

seek to overcome the odds of RAEs by competing with and challenging their older and 'more advanced' peers [11, 23]. As such, previous research has sought to measure players' long-term development via the examination of the wages of professional German football players [2] and the estimated market value of U18–U23 professional football players [30], in relation to birthdate distributions. Findings from these studies provide support for the underdog hypothesis, in which those born later in the year (i.e. Q4) earned systematically higher wages [2] and an undervaluing of Q4 players in younger age groups, with an increase in their market value over time. An examination of monetary values (e.g. estimated market value) provide a viable means by which the (perceived) value of professional players and their long-term development can be examined. This approach can be extended to other European leagues, including England, thus providing an improved insight and understanding into the extent of the proposed 'underdog hypothesis'.

The aim of this research was to examine the prevalence of RAEs within U18, U21 and professional senior squads that compete within the highest (respective) leagues within England. Moreover, this research sought to examine RAEs in relation to position (GK, Def, Mid and Att), age (between age groups and across age bands within each age group), and estimated market value within senior players. In doing so this study makes a significant and original contribution to the existing RAEs research, providing an overview of the extent to which the birthdate distribution of U18 and U21 squads is representative of the birthdate distribution within professional senior squads that compete within the highest (respective) leagues within England. Continued examination and analysis of RAEs acts as a useful resource for those working within youth football academies, aiding in the development and promotion of an improved understanding and recognition of these RAEs in U18, U21 and senior squads, particularly when compared to senior professional squads. Moreover, the innovative use of estimated market value as a proxy for long-term development provides a novel means by which the existence of the underdog hypothesis can be monitored. This may then help to minimise the prevalence of any selection biases within younger age groups, whilst also supporting and facilitating the implementation of potential solutions (to combat and minimise the potential effects of RAEs) to aid (long-term) talent development processes.

Methods

Consistent with the approach taken by Romann et al. [30], data were obtained from publicly available sources (https://www.transfermarkt.co.uk/) for squads competing in the U18s Premier League North (13 clubs, $n = 271$) and South

(12 clubs, $n = 216$) leagues, Premier League 2 (14 clubs, $n = 350$) and the corresponding senior squads, who competed in either the Premier League (13 clubs, $n = 371$) or Championship (1 club, $n = 25$) during the 2022–23 season. All data related to 2022–23 season and were obtained at the end of the season in June 2023. For each player, Club, Date of Birth, Birth Month, Playing Position (Goalkeeper, Defender, Midfielder or Attacker), Birth Quartile and Age (years.) were recorded. In addition, whether players were 'UK' nationals or 'Non-UK' nationals were also recorded for players competing in Premier League 2 (U21 players) and Premier League senior squads, with estimated market value also recorded for senior players. All relevant data were extracted and inputted into specially designed Excel spreadsheets for further analysis. Individuals' birthdates were categorised into relative birth quartiles (BQ) in accordance with the selection year (Q1 = Sept–Nov; Q2 = Dec–Feb; Q3 = Mar–May; Q4 = Jun–Aug), which is employed throughout competitive youth football within England. Ethical approval was granted by the respective University's Departmental Research and Ethics Committee (ETH2223-0349).

Data Analysis

Following procedures outlined by McHugh [22], Chi-square (χ^2) analysis was employed to compare quartile distributions in the sample and against population values [24] and expected birthdate distributions for within (youngest age band within each squad) and across (U18 data) age group analysis. As the Chi-square test does not reveal the magnitude of difference between quartile distributions for significant chi square outputs, Cramer's V and Odds Ratios (OR) with 95% confidence intervals (CI) were also calculated to examine the bias of birth-date distributions within groups (Q1, Q2, Q3 and Q4). The Cramer's V was interpreted as per conventional thresholds (i.e., ≥ 0.06 = small effect size, ≥ 0.17 = medium effect size, and ≥ 0.29 = large effect size) [8]. The OR was calculated to compare the birth-date distribution of a particular quartile (Q1, Q2 or Q3) with the reference group, which consisted of the relatively youngest players (Q4) for the representative group. A higher OR indicates an increased representation of players who were born in that quartile compared to the reference quartile Q4. These will be considered significant when the CI range did not include a value ≤ 1.00. Finally, where appropriate the alpha level was set at $P < 0.05$.

Results

The frequency and percentage distributions of players' birth quartiles within each age group are presented in Table 1. In addition, Table 1 also provides the frequency and percentage

Table 1 Birthdate distribution and analysis across and within U18, U21 and Senior squads

Competition	Age	n	Birthdate Distribution n(%) Q1	Q2	Q3	Q4	Odds Ratios (95% CI) Q1 vs. Q4	Q2 vs. Q4	Q3 vs. Q4	Chi² X^2_3	P Value	V
U18	≤17 years	66	17 (25.8)	13 (19.7)	4 (6.1)	32 (48.5)	0.5 (0.3–0.9)	0.4 (0.2–0.7)	0.1 (0.07–0.2)	35.6*	<0.001	0.42
	17–18 years	248	116 (46.8)	64 (25.8)	41 (16.5)	27 (10.9)	4.3 (2.5–7.5)	2.4 (1.4–4.1)	1.5 (0.9–2.6)	29.5*	<0.001	0.20
	≥18 years	173	79 (45.7)	46 (26.6)	34 (19.7)	14 (8.1)	5.6 (3.2–9.8)	3.3 (1.9–5.7)	2.4 (1.4–4.2)	29.5*	<0.001	0.24
	All	487	212 (43.5)	123 (25.3)	79 (16.2)	73 (15.0)	2.9 (1.7–5.1)	1.7 (1.0–2.9)	1.1 (0.6–1.9)	20.4*	<0.001	0.12
Premier League 2	≤19 years	96	41 (42.7)	25 (26.0)	12 (12.5)	18 (18.8)	2.3 (1.3–4.0)	1.4 (0.8–2.4)	0.7 (0.4–1.2)	19.9	<0.001	0.26
	19–20 years	115	63 (54.8)	23 (20.0)	20 (17.4)	9 (7.8)	7.0 (4.0–12.2)	2.6 (1.5–4.5)	2.2 (1.3–3.9)	50.2	<0.001	0.38
	≥20 years	139	57 (41.0)	36 (25.9)	34 (24.5)	12 (8.6)	4.8 (2.7–8.3)	3.0 (1.7–5.2)	2.8 (1.6–4.9)	21.1	<0.001	0.22
	All	350	161 (46.0)	84 (24.0)	66 (18.9)	39 (11.1)	4.1 (2.4–7.2)	2.2 (1.2–3.8)	1.7 (1.0–3.0)	26.5	<0.001	0.16
Senior Squads	≤24 years	138	40 (29.0)	36 (26.1)	31 (22.5)	31 (22.5)	1.3 (0.7–2.3)	1.2 (0.7–2.0)	1.0 (0.6–1.7)	1.2	=0.74	0.05
	24–28 years	133	35 (26.3)	39 (29.3)	32 (24.1)	27 (20.3)	1.3 (0.7–2.3)	1.4 (0.8–2.5)	1.2 (0.7–2.1)	2.2	=0.54	0.07
	≥28years	125	26 (20.8)	32 (25.6)	37 (29.6)	30 (24.0)	0.9 (0.5–1.5)	1.1 (0.6–1.9)	1.2 (0.7–2.1)	1.9	=0.59	0.07
	All	396	106 (26.8)	102 (25.8)	101 (25.5)	87 (22.0)	1.2 (0.7–2.1)	1.2 (0.7–2.0)	1.2 (0.7–2.0)	0.7	=0.88	0.02

Q1 Sep–Nov, *Q2* Dec–Feb, *Q3* Mar–May, *Q4* Jun–Aug

*Significant effect at an alpha level of $P < 0.05$

distributions of birth quartiles for distinct age bands within each cohort. The Chi-square test showed significant deviations across birth quartiles for U18 [χ^2 (df = 3) = 20.4, $P < 0.001$, $V = 0.12$] and U21 [χ^2 (df = 3) = 26.5, $P < 0.001$, $V = 0.16$] age groups, and for all age bands within the U18 and U21 age groups. However, the significant deviation [χ^2 (df = 3) = 35.6, $P < 0.001$, $V = 0.42$] for those aged ≤ 17 years old, within the U18 age group, was due to an over-representation of players born within Q4 (48.5%), which is supported by the OR analysis. Indeed, in both U18 and U21 age groups analysis revealed a greater proportion of relatively younger players within the youngest age bands (≤ 17 years old & ≤ 19 years old, respectively). The RAEs were more prevalent within the older age bands in both the U18 (≥ 18 years old; [χ^2 (df = 3) = 29.5, $P < 0.001$, $V = 0.24$] and U21 [≥ 20 years old; χ^2 (df = 3) = 21.1, $P < 0.001$, $V = 0.22$] age groups, with a stratified over-representation of those born earlier within the selection year clearly displayed (i.e. Q1 > Q2 > Q3 > Q4). In further support of this, the frequency of birthdate distributions according to birth month are displayed in Fig. 1, providing additional evidence of RAEs within U18 (Fig. 1a) and U21 (Fig. 1b) age groups, for

players competing at the highest level of competition within their respective ages. In contrast, there were no significant deviations across birth quartiles within senior squads [χ^2 (df = 3) = 4.0, $P = 0.26$, $V = 0.06$], which was also supported by the lower OR values, demonstrating a more equal birthdate distribution (Fig. 1c).

Analysis comparing birthdate distributions between age groups, where expected birthdate distributions were based on U18 data and compared to birthdate distributions of senior squads and U21 squads, revealed no significant difference between U18 and U21 squads [χ^2 (df = 3) = 1.67, $P = 0.64$, $V = 0.04$] but a significant difference between U18 and senior squads [χ^2 (df = 3) = 15.0, $P < 0.001$, $V = 0.11$]. Similarly, there was a significant difference between U21 and senior squads [χ^2 (df = 3) = 21.2, $P < 0.001$, $V = 0.13$] when expected birthdate distributions were based on U21 data. Moreover, analysis of birthdate distributions across age bands within each age group, where expected birthdate distributions were based on the youngest age band within each age group, revealed a significant difference between ≤ 17 years old and 17–18 years old [χ^2 (df = 3) = 65.9, $P < 0.001$, $V = 0.30$],
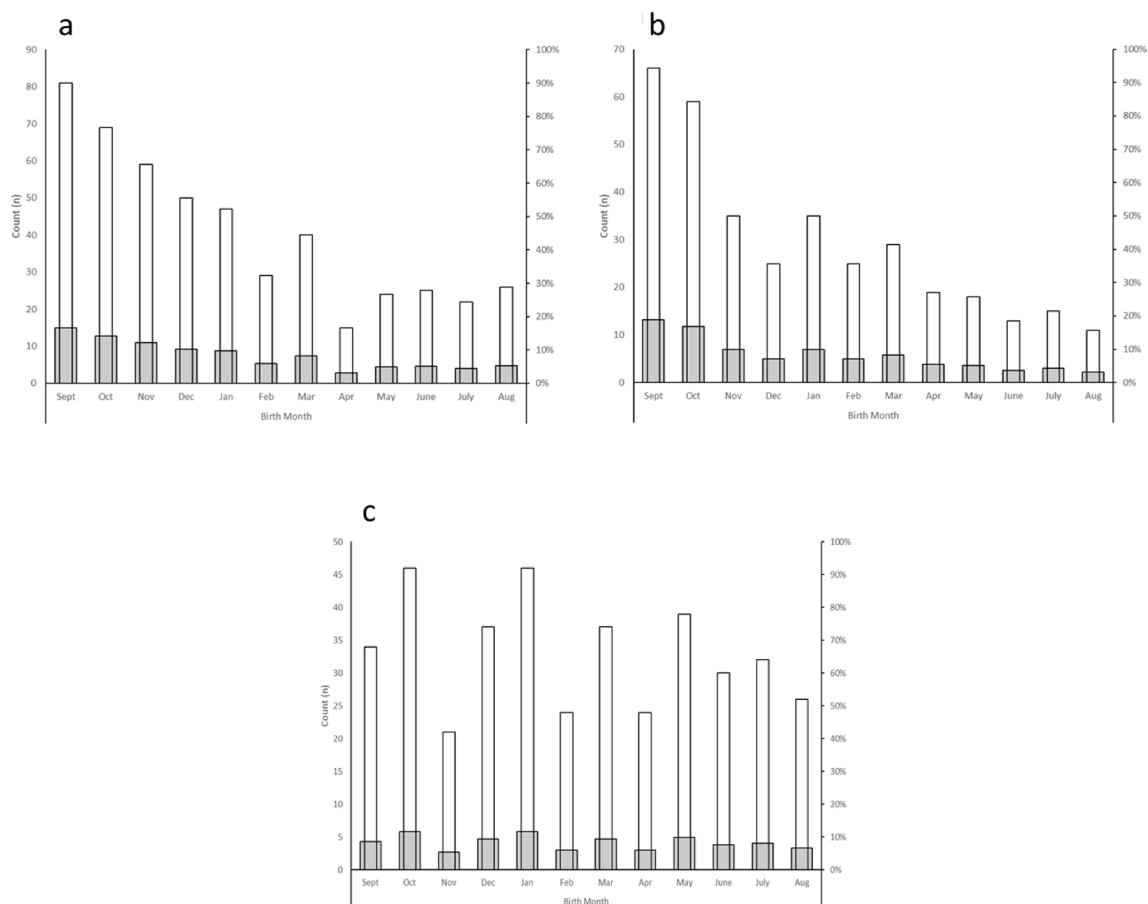

Fig. 1 The frequency of birthdate distributions according to birth month for U18 (**a**), U21 (**b**) and Senior (**c**) age groups

and between ≤ 17 years old and those aged > 18 years old [χ^2 (df = 3) = 81.7, $P < 0.001$, $V = 0.40$] in the U18 squads. Analysis of age bands within U21 squads also revealed a significant difference between ≤ 19 years old and 19–20 years old χ^2 (df = 3) = 13.2, $P < 0.001$, $V = 0.20$], as well as between ≤ 19 years old nd > 20 years old [χ^2 (df = 3) = 17.1, $P < 0.001$, $V = 0.20$]. There were no significant differences in birthdate distributions between age bands within senior squads.

The frequency and percentage distributions of players' birth quartiles within each age group according to position are presented in Table 2. The Chi-square test showed significant deviations across birth quartiles for U18 and U21 age groups, and for all positions (GK, Def, Mid and Att) within the U18 and U21 age groups ($P \leq 0.01$). Moreover, OR analysis revealed a distinct over-representation of players born in Q1 for all positions within U18 and U21 age groups. In contrast, there were no significant deviations across birth quartiles according to position within senior squads. Although, OR analysis still suggested a small (but non-significant) number of players born in Q1, Q2 and Q3, in comparison to those born in Q4 in senior squads. Analysis also revealed a substantial change in the relative contribution of 'UK' nationals and 'Non-UK' nationals to teams within U21 (277 [79.1%] UK nationals and 73 [20.9%] Non-UK nationals) and senior (167 [42.2%] UK nationals and 229 [57.8%] Non-UK nationals) squads. Despite the increased contribution of 'Non-UK' nationals within senior squads, there were no significant deviations across birth quartiles for both UK [χ^2 (df = 3) = 3.3, $P = 0.35$, $V = 0.08$] and Non-UK [χ^2 (df = 3) = 1.2, $P = 0.75$, $V = 0.04$] nationals (Fig. 2). Finally, although not statistically significant, analysis of senior players' estimated market value revealed a higher average estimated market value for players born in Q4 in comparison to those born in Q1, Q2 and Q3 (Fig. 3).

Discussion

The current study sought to examine the prevalence of RAEs within U18, U21 and professional adult squads that compete within the highest (respective) leagues within England, as well as considering playing position (GK, Def, Mid and Att) and market value. Analysis revealed significant RAEs within U18 and U21 age groups, whereby more relatively older players were selected than their relatively younger counterparts. Further analysis of age bands within each age group also revealed an increase in the prevalence of RAEs throughout each age group, with the youngest age bands within the U18 (≤ 17 years) and U21 (≤ 19 years) displaying smaller OR than those reported for the older age bands within each age group. This selection bias was evident across all positions (GK, Def, Mid and Att) within U18 and U21 age groups. In contrast, analysis of senior squads revealed no significant deviations in birthdate distributions when considered as a whole sample, as separate age bands or by position. However, analysis of senior players' estimated market value found the average estimated market value of Q4 players to be the highest (Fig. 3), providing possible support for the proposed 'underdog hypothesis'.

The reported bias towards selecting players born early in the selection year and subsequent over-representation of relatively older players, within younger (U18 and U21) age groups, is consistent with the findings of previous research within youth male football [1, 3, 18, 21, 23]. As an example, Andrew et al. [1] reported significant RAEs within U17 and

Table 2 Birthdate distribution and analysis across positions, within U18, U21 and Senior squads

Competition	Age	n	Birthdate Distribution n(%)				Odds Ratios (95% CI)			Chi^2	P Value	V
			Q1	Q2	Q3	Q4	Q1 vs. Q4	Q2 vs. Q4	Q3 vs. Q4	X^2_3		
U18	GK	51	28 (54.9)	15 (29.4)	3 (5.9)	5 (9.8)	5.6 (3.2–9.8)	3.0 (1.7–5.2)	0.6 (0.3–1.0)	59.7*	< 0.001	0.48
	Def	160	69 (43.1)	44 (27.5)	22 (13.8)	25 (15.6)	2.8 (1.6–4.8)	1.8 (1.0–3.1)	0.9 (0.5–1.5)	21.6*	< 0.001	0.25
	Mid	156	70 (44.9)	31 (19.9)	32 (20.5)	23 (14.7)	3.0 (1.8–5.3)	1.4 (0.8–2.4)	1.4 (0.8–2.4)	21.3*	< 0.001	0.25
	Att	120	45 (37.5)	33 (27.5)	22 (18.3)	20 (16.7)	2.3 (1.3–3.9)	1.7 (1.0–2.9)	1.1 (0.6–1.9)	11.0*	= 0.01	0.24
Premier League 2	GK	46	20 (43.5)	11 (23.9)	8 (17.4)	7 (15.2)	2.9 (1.6–5.0)	1.6 (0.9–2.7)	1.1 (0.7–2.0)	23.4*	< 0.001	0.38
	Def	118	48 (40.7)	37 (31.4)	21 (17.8)	12 (10.2)	4.0 (2.3–7.0)	3.1 (1.8–5.4)	1.8 (1.0–3.1)	22.6*	< 0.001	0.38
	Mid	103	53 (51.5)	16 (15.5)	22 (21.4)	12 (11.7)	4.4 (2.5–7.7)	1.3 (0.8–2.3)	1.8 (1.1–3.2)	38.2*	< 0.001	0.26
	Att	83	40 (48.2)	20 (24.1)	15 (18.1)	8 (9.6)	5.0 (2.9–8.7)	2.5 (1.4–4.4)	1.9 (1.1–3.3)	32.4*	< 0.001	0.33
Senior Squads	GK	48	12 (25.0)	10 (20.8)	17 (35.4)	9 (18.8)	1.3 (0.8–2.3)	1.1 (0.6–1.9)	1.9 (1.1–3.3)	6.8	= 0.08	0.09
	Def	137	38 (27.7)	35 (25.5)	37 (27)	27 (19.7)	1.4 (0.8–2.5)	1.3 (0.7–2.3)	1.4 (0.8–2.4)	1.8	= 0.61	0.07
	Mid	111	31 (27.9)	32 (28.8)	23 (20.7)	25 (22.5)	1.2 (0.7–2.2)	1.3 (0.7–2.2)	0.9 (0.5–1.6)	2.1	= 0.55	0.08
	Att	100	25 (25.0)	25 (25.0)	24 (24.0)	26 (26.0)	1.0 (0.6–1.7)	1.0 (0.6–1.7)	0.9 (0.5–1.6)	0.1	= 1.00	0.05

Q1 Sep–Nov, *Q2* Dec–Feb, *Q3* Mar–May, *Q4* Jun–Aug

*Significant effect at an alpha level of $P < 0.05$

Fig. 2 Birthdate distribution across quartiles for UK and Non-UK nationals, within Senior squads
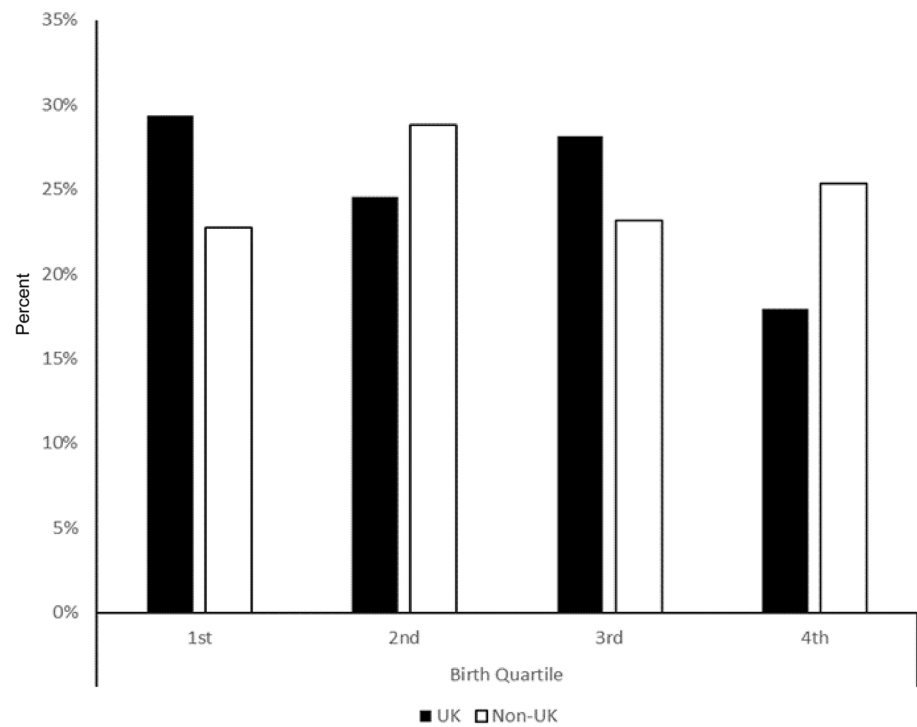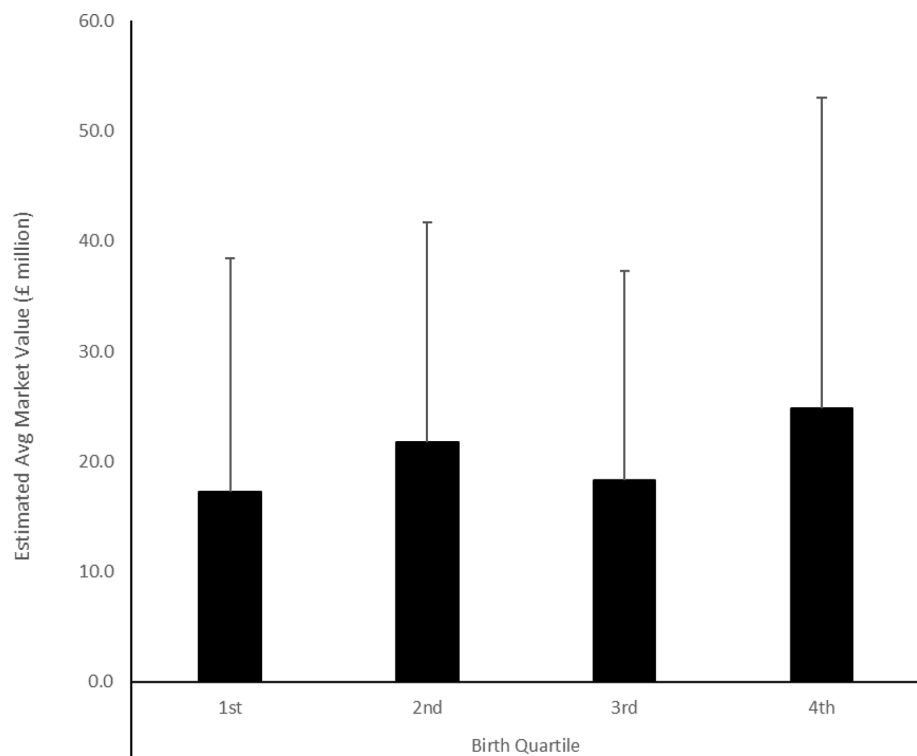

Fig. 3 Estimated average market value (per million £) across birth quartiles, within Senior squads

U19 male European football players that competed in European Championship qualification campaigns, but not at senior level. Moreover, Andrew et al. [1] found that RAEs were more prevalent within male U17 and U19 age groups for those that played for teams which qualified, in comparison to those who did not qualify. This supports previous research that suggests there is a possible link between the level of 'competition' and the prevalence of the RAEs, whereby increased levels of competition (i.e., competing for selection) results in larger RAEs and a greater over-representation

of players born in earlier in the selection year [5, 7, 10, 27]. As the current findings are from those competing at the highest levels of competitive football in England, within their respective age groups, the enhanced levels of 'competitiveness' may exacerbate the prevalence of RAEs within male youth football. Indeed, the increased competition for selection as players progress throughout each age group (i.e. age bands within each age group) appears to result in an increased prevalence of the RAEs, as evidenced by the increased OR between quartiles in the older age bands, within the U18 and U21 age groups. This prompts questions regarding the processes and procedures in relation to talent identification and talent development within high-level male youth football, not only across age groups but within age groups too, which is an area that has received limited attention within the literature.

Findings also revealed that RAEs were prevalent in both U18 and U21 age groups in accordance with on-field position (GK, Def, Mid and Att), corresponding with previous research in youth male football [28, 32, 33]. Specifically, current findings show that, within the U18 age group, RAEs were most pronounced for goalkeepers, whereas it was greatest for attackers within the U21 age group. Moreover, in contrast to goalkeepers, all out-field positions (Def, Mid and Att) revealed more pronounced RAEs in U21 age groups, in comparison to their U18 counterparts. The over-representation of attackers within the U21 age group is supported by Peña-González et al. [28], who also reported an over-representation of relatively older players in attacking positions within similar age groups, at the U19/U21 Championship, South American U20 Championship, and U20 World Cup in 2019. In contrast, recent research by McAuley et al. [21] reported that a greater number of relatively younger attackers were selected, when examining RAEs across playing levels and positions in Northern Ireland (recognised as an emerging nation) international male football. This provides further support for the complex and multifaceted nature of RAEs [7, 34], and the extent to which RAEs appear to be influenced by various socio-cultural and contextual factors. Indeed, research by Doncaster et al. [10] provisionally hypothesized that based on FC Barcelona's talent identification model, RAEs would be less prevalent than has been previously reported within the literature. Yet, results found that RAEs were prevalent throughout all male football age groups, from U10 to senior levels [10]. This contrasts with current findings, which show a less pronounced (non-significant) RAEs at senior level. Consequently, future research should seek to provide greater recognition and attention to the broader socio-cultural and contextual factors in relation to RAEs. Here, the proposed theoretical models of Hancock et al. [12] and Wattie et al. [34] to better explain RAEs in sport should be investigated and applied to empirical studies to a greater extent.

In contrast to the U18 and U21 age groups, and in line with previous research [1, 21], RAEs were less pronounced at senior level. However, while McAuley et al. [21] revealed a progressive decline in the prevalence of RAEs from U17 through to senior level within Northern Irish international male football players, present findings display an increase in the prevalence of RAEs between U18 and U21 age groups, before substantially reduced RAEs at senior level. The specific mechanisms underpinning the variations in reported RAEs across the literature are unknown but are likely a combination of various factors relating to the physical characteristics, specific sociocultural context and playing styles across (and within) clubs. In addition, the present study offers consideration to the national status (UK or Non-UK) of players, demonstrating a substantial increase in the relative contribution of Non-UK players from U21 to senior squads (20.9%–57.8%, respectively). Given the differences in the age group selection cut-off dates between UK (Sept–Aug) and Non-UK (Jan–Dec) it could be argued that the diminished RAEs at senior level within the current study is a result of the increased relative contribution of Non-UK players. Further analysis of present data, however, found that Non-UK players within senior age groups, utilising January to December cut-off dates, corroborated results that demonstrated diminished RAEs at senior level (1st Quartile $n = 59$ [25.8%], 2nd Quartile $n = 58$ [25.3%], 3rd Quartile $n = 54$ [23.6%], 4th Quartile $n = 58$ [25.3%]).

The higher average estimated market value for senior players who are born later in the selection year (Q4), provides further support for the underdog hypothesis [11, 17] and adds to the existing literature which has explored estimated market value as a proxy for success within senior football [30]. In support of the present results, Romann et al. [30] reported an undervaluing of Q4 players in younger age groups, but then an increase in estimated market value as players aged. In contrast, players born earlier in the selection year (Q1) suggested an initial over-estimation of market value within younger age groups, which then decreased as players aged, and progressed into senior level football. Previously, research has proposed that the reduced prevalence of RAEs in older age groups (i.e., senior level) may be explained by the underdog hypothesis [9, 11, 17]. The underdog hypothesis proposes that the advantages associated with an older relative age become attenuated near adulthood, perhaps because the improved psychological, social, technical and tactical skills developed by relatively younger players during adolescence become more salient [11]. In addition, relatively younger individuals who remain in the talent development system may experience a comparatively greater challenge than their relatively older peers, which could facilitate the improvement of psychological, social, technical and tactical skills, manifesting in superior performance capabilities within adulthood. Finally, at older ages,

the difference between relatively older and younger players in terms of experience and opportunities to practice are also reduced (i.e., an 11-month difference in age at 10-years-old represents a 10% difference, whereas at 20-years-old an 11-month difference only represents a 5% difference).

In support of the multifaceted and contextual nature of RAEs it may be argued that the occurrence and impact of the underdog hypothesis is similarly affected by contextual factors. Indeed, as in the present study, the highest levels of competition (with large talent pools) may result in a delayed impact of the underdog hypothesis, resulting in a prolonged (and potentially increased) prevalence of RAEs. Whereas lower levels of competition (with smaller talent pools) may result in an early onset and influence of the underdog hypothesis, reducing the prevalence of the RAEs from a younger age. As such, the use and application of longer-term outcome measures (i.e., estimated market value, appearances and wages) in conjunction with birth-date distribution data may provide an improved understanding of the extent to which RAEs within senior football are representative of youth football.

As noted by McAuley et al. [21], it is important athlete development systems and pathways are organised with an equitable framework to ensure they: (a) are efficient and effective, (b) reduce talent wastage by promoting talent inclusion, and (c) prioritise future potential over current performance [4]. To date, however, particularly within high-performance youth male football environments, the processes and strategies commonly employed are unable to mitigate factors that confound the identification, development, and selection processes in youth male football, thus resulting in the continued identification of RAEs [1, 3, 18, 23]. This is despite the instigation of the Premier League's EPPP [29] and the formal recognition (and subsequent research) of RAEs within youth male football. Whilst there is a growing appreciation of the varying mechanisms that affect the prevalence of RAEs, including, physical (e.g., anthropometric and physiological characteristics), psychological and sociocultural, continued research is needed to acknowledge the ever-changing and specific contexts across high-performance youth football settings. Here, there should be continued efforts to monitor and track the prevalence and persistence of RAEs throughout the talent development pathway within youth male football, but also improved attempts to investigate, analyse and evaluate the various processes involved in talent identification processes in which explanations and justifications for (de)selection by key stakeholders (e.g., coaches, scouts, sports practitioners) are sought. Indeed, like Ludin et al. [19], future research should seek to better understand the talent identification and selection processes employed by key stakeholders and whether an appropriate level of appreciation is given to players' relative age and long-term potential. Here, education for practitioners regarding RAEs and the associated implications for player development and talent identification should also be considered.

It is recognised that the absence of individual teams and performance outcomes (e.g. league position) are a limitation of the current study. Such data would be beneficial as it would enable the prevalence of RAEs to be analysed alongside performance-based metrics. In this regard, however, the purpose of the current study was to investigate the prevalence of relative age effects (RAEs) within and between U18, U21 and professional senior squads, that compete in the highest (respective) leagues within England, irrespective of performance outcomes. Nevertheless, future research should seek to consider the RAEs alongside various performance metrics. Furthermore, whilst the current study provides an indication of the RAEs within and across age groups, the cross-sectional design resulted in an inability to accurately assess the prevalence of RAEs between age groups (i.e. the key transitional periods from one age group to another). As such, future research should seek to adopt a longitudinal research design in which cohorts are tracked over numerous seasons and age groups (e.g. U14 through to senior football). Here, consideration should also be given towards the retained, released (including their destinations) and recruited players (including their previous club) in respect of the RAEs. Ultimately, an improved tracking of the movement of players during their development will provide greater insights into the underdog hypothesis. Finally, future research should be conducted within and across various European domestic leagues, thus allowing for an improved understanding of the RAEs during key developmental stages within highly trained youth football players.

Practical Implications

Despite the growing acknowledgement, wide array of research and formal recognition (by the Premier League) of RAEs, further work is still required to help combat the prevalence of RAEs and provide organised and equitable athlete identification systems and development pathways. Indeed, if not already available, the provision of a national database which provides an ongoing record and analysis of player birthdate-distributions within and across (EPPP) youth football academies should be feasible. The development of such a database would require clubs to provide the Premier League with an up-to-date record of players registered (signed) within their academy, which can be anonymised to ensure only the relative birthdate distribution of players (i.e., percentage of players born in September) is accessible. This would allow for the longitudinal tracking of RAEs and the opportunity for further research to be undertaken in relation to age groups, positions as well as (de)selection of players. Furthermore, a national database would likely provide an

improved recognition and understanding of RAEs, in relation to key performance indicators, thus supplementing and complimenting coach development programmes.

To aid in-competition talent identification processes, the implementation of age-ordered shirt numbers should be considered [20]. This should support talent scouts and identification processes with the knowledge that the numbers on the playing shirts corresponded with the relative age of the players, helping those involved in the selection process to consciously consider current and potential talent of players in consideration of their relative age within the players being observed. Consistent with previous research, however, RAEs are not the only variable to consider in practice, and other variables such as growth and maturation need to be accounted for and appreciated within the complex operations of talent identification and development. Helsen et al. [14] propose a novel method in which players' maturity-status and relative age are both considered to minimise or nullify inequalities resulting from relative-age and maturity-related biases. Following the reallocation of players, Helsen et al. [14] reported a more even distribution of birthdates (i.e. a reduction in RAEs) throughout a selection year, alongside a reduction in stature and body mass differences within a 'reallocated' cohort.

Conclusion

This study supplements the existing literature exploring RAEs by providing original insights regarding the prevalence of RAEs in U18, U21 and senior squads that compete in the highest (respective) leagues within England, in consideration of position and age. Moreover, the novel application and analysis of senior players' estimated market value offers an innovative means by which the existence and extent of the proposed underdog hypothesis could be examined. This, in turn, may provide useful evidence concerning the long-term development of players and an improved appreciation for long-term potential rather than current performance. Nevertheless, RAEs continue to manifest across academies in England, this is despite the ongoing research to highlight its prevalence from stakeholders (i.e., researchers, policy makers, practitioners), including calls to action from its own governing body (Premier League [29]). This suggests that despite the increasing awareness of the existence of RAEs in academies, practitioners continue to engage in activities that are inherently biased. A greater appreciation of practitioners knowledge, understanding and practices in relation to RAEs are required to inform future educational strategies. Therefore, those responsible for the organisational structures should place an emphasis on strategies to moderate RAEs, whilst also considering strategies to stretch and challenge all young players (e.g., 'playing-up') [16]. Further

consideration toward the design, implementation, and evaluation of these relative age strategies is required, ensuring that they support the development of every young player to maximise their long-term development.

Declarations

References

1. Andrew M, Finnegan L, Datson N, Dugdale JH. Men are from quartile one, women are from? Relative age effect in European soccer and the influence of age, success, and playing status. Children. 2022;9(11):1747. https://doi.org/10.3390/children9111747.
2. Ashworth J, Heyndels B. Selection bias and peer effects in team sports: the effect of age grouping on earnings of German soccer players. J Sports Econ. 2007;8(4):355–77. https://doi.org/10.1177/1527002506287695.
3. Augste C, Lames M. The relative age effect and success in German elite U-17 football teams. J Sports Sci. 2011;29(9):983–7. https://doi.org/10.1080/02640414.2011.574719.
4. Baker J, Schorer J, Wattie N. Compromising talent: issues in identifying and selecting talent in sport. Quest. 2018;70(1):48–63. https://doi.org/10.1080/00336297.2017.1333438.
5. Bezuglov EN, Nikolaidis PT, Khaitin V, Usmanova E, Luibushkina A, Repetiuk A, Knechtle B. Prevalence of relative age effects in Russian soccer: the role of chronological age and performance. Int J Environ Res Public Health. 2019. https://doi.org/10.3390/ijerph16214055.
6. Castillo D, Pérez-González B, Raya-González J, Fernández-Luna Á, Burillo P, Lago-Rodríguez Á. Selection and promotion processes are not associated by the relative age effect in an elite Spanish soccer academy. PLoS ONE. 2019;14(7): e0219945. https://doi.org/10.1371/journal.pone.0219945.
7. Cobley S, Baker J, Wattie N, McKenna J. Annual age-grouping and athlete development: a meta-analytical review of relative age effects in sport. Sports Med. 2009;39(3):235–56.

8. Cohen J. Statistical power analysis for the behavioral sciences. Hillsdale (NJ): L. Erlbaum Associates; 1988.

9. Cumming SP, Searle C, Hemsley JK, Haswell F, Edwards H, Scott S, Malina RM. Biological maturation, relative age and self-regulation in male professional academy soccer players: a test of the underdog hypothesis. Psychol Sport Exercise. 2018;39:147–53.

10. Doncaster G, Medina D, Drobnic F, Gómez-Díaz AJ, Unnithan V. Appreciating factors beyond the physical in talent identification and development: insights from the FC Barcelona sporting model. Front Sports Act Living. 2020;2:91.

11. Gibbs BG, Jarvis JA, Dufur MJ. The rise of the underdog? The relative age effect reversal among Canadian-born NHL hockey players: a reply to Nolan and Howell. Int Rev Sociol Sport. 2012;47(5):644–9.

12. Hancock D, Adler A, Cote J. A proposed theoretical model to explain relative age effects in sport. Eur J Sports Sci. 2013;13(6):630–7. https://doi.org/10.1080/17461391.2013.775352.

13. Helsen WF, Baker J, Michiels S, Schorer J, Van Winckel J, Williams AM. The relative age effect in European professional soccer: did ten years of research make any difference? J Sports Sci. 2012;30(15):1665–71. https://doi.org/10.1080/02640414.2012.721929.

14. Helsen WF, Thomis M, Starkes JL, Vrijens S, Ooms G, MacMaster C, Towlson C. Levelling the playing field: a new proposed method to address relative age-and maturity-related bias in soccer. Front Sports Act Living. 2021;3:24.

15 Kelly A. Talent identification and development in youth soccer: a guide for researchers and practitioners. 1st ed. New York: Routledge; 2023. https://doi.org/10.4324/9781032232799.

16. Kelly AL, Wilson MR, Jackson DT, Goldman DE, Turnnidge J, Côté J, Williams CA. A multidisciplinary investigation into "playing-up" a chronological age group in an English football academy. J Sports Sci. 2021;39(8):854–64. https://doi.org/10.1080/02640414.2020.1848117.

17 Kelly AL, Wilson MR, Gough LA, Knapman H, Morgan P, Cole M, Williams CA. A longitudinal investigation into the relative age effect in an English professional football club: exploring the 'underdog hypothesis'. Sci Med Football. 2020;4(2):111–8.

18. Lovell R, Towlson C, Parkin G, Portas M, Vaeyens R, Cobley S. Football player characteristics in English lower-league development programmes: the relationships between relative age, maturation, anthropometry and physical fitness. PLOSone. 2015;10(9):e0137238. https://doi.org/10.1371/journal.pone.0137238.

19. Lüdin D, Donath L, Romann M. Disagreement between talent scouts Implications for improved talent assessment in youth football. J Sports Sci. 2023;41(8):758–65. https://doi.org/10.1080/02640414.2023.2239614.

20. Mann DL, van Ginneken PJMA. Age-ordered shirt numbering reduces the selection bias associated with the relative age effect. J Sports Sci. 2017;35(8):784–90.

21. McAuley ABT, Baker J, Johnston K, Doncaster G, Kelly AL. Selection and re-selection throughout a national talent pathway: exploring longitudinal relative age effects in Northern Ireland male soccer Research for Athlete and Youth Sport Development (RAYSD) lab, Birmingham City University. 2024. https://doi.org/10.3390/sports9070101

22. McHugh ML. The chi-square test of independence. Biochemia Med. 2013;23(2):143–9.

23. Morganti G, Kelly AL, Apollaro G, Pantanella L, Esposito M, Grossi A, Ruscello B. Relative age effects and the youth-to-senior transition in Italian football: the underdog hypothesis versus knock-on effects of relative age. Sci Med Football. 2022;7(4):406–12.

24. Office for National Statistics. Number of live births by date, 1995 to 2014, in England and Wales. 2015. https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/livebirths/articles/howpopularisyourbirthday/2015-12-18#:~:text=September%2026%20was%20the%20most,other%202%20being%20early%20October. Accessed 20 July 2023.

25. Padrón-Cabo A, Rey E, García-Soidán JL, Penedo-Jamardo E. Large scale analysis of relative age effect on professional soccer players in FIFA designated zones. Int J Perform Anal Sport. 2016;16(1):332–46. https://doi.org/10.1080/24748668.2016.11868890.

26. Pedersen AV, Aune TK, Dalen T, Lorås H. Variations in the relative age effect with age and sex, and over time-Elite-level data from international soccer world cups. PLoS ONE. 2022;17(4):e0264813. https://doi.org/10.1371/journal.pone.0264813.

27. Pena-Gonzalez I, Fernandez-Fernandez J, Moya-Ramon M, Cervello E. Relative age effect, biological maturation, and coaches' efficacy expectations in young male football players. Res Q Exerc Sport. 2018;89(3):373–9.

28. Peña-González I, Javaloyes A, Sarabia JM, Moya-Ramón M. Relative age-related differences between different competitive levels and field positions in young soccer players. Res Sports Med. 2021;29(3):254–64. https://doi.org/10.1080/15438627.2020.1853540.

29. Premier League. Elite Player Performance Plan. 2011. https://www.goalreports.com/EPLPlan.pdf. Accessed 20 July 2023.

30. Romann M, Javet M, Cobley S, Born D. How relative age effects associate with football players' market values: Indicators of losing talent and wasting money. Sports. 2021;9(7):99. https://doi.org/10.3390/sports9070099.

31. Unnithan V, White J, Georgiou A, Iga J, Drust B. Talent identification in youth soccer. J Sports Sci. 2012;30(15):1719–26. https://doi.org/10.1080/02640414.2012.731515.

32 Salinero JJ, Pérez B, Burillo P, Lesma ML. Relative age effect in European professional football: analysis by position. J Hum Sport Exercise. 2013;8(4):966–73. https://doi.org/10.4100/jhse.2013.84.07.

33. Towlson C, Cobley S, Midgley AW, Garrett A, Parkin G, Lovell R. Relative age, maturation and physical biases on position allocation in elite-youth soccer. Int J Sports Med. 2017;38(3):201–9. https://doi.org/10.1055/s-0042-119029.

34. Wattie N, Schorer J, Baker J. The relative age effect in sport: a developmental systems model. Sports Med. 2015;45(1):83–94. https://doi.org/10.1007/s40279-014-0248-9.

35. Yagüe JM, de la Rubia A, Sánchez-Molina J, Maroto-Izquierdo S, Molinero O. The relative age effect in the 10 best leagues of male professional football of the Union of European Football Associations (UEFA). J Sports Sci Med. 2018;17(3):409–16.