The prevalence of relative age effects and the influence of the talent pool size on Australian male and female youth football

Kyle J.M. Bennett [a*], Andrew R. Novak [b], Job Fransen [b], and Rob Duffield [c,d]

[a] *Discipline of Sport and Exercise Science, Faculty of Health, Southern Cross University, Coffs Harbour, Australia.* [b] *Human Performance Research Centre, School of Sport, Exercise, and Rehabilitation, Faculty of Health, University of Technology Sydney, Moore Park, Australia,* [c] *School of Sport, Exercise, and Rehabilitation, Faculty of Health, University of Technology Sydney, Moore Park, Australia.* [d] *Medical Department, Football Australia, Sydney, Australia.*

*Corresponding author:

Kyle J.M. Bennett
Discipline of Sport and Exercise Science
Faculty of Health
Southern Cross University
Hogbin Drive, Coffs Harbour NSW 2452
Kyle.Bennett@scu.edu.au

ORCID

Kyle Bennett (0000-0002-2288-4770), Andrew Novak (0000-0002-2949-4150), Job Fransen (0000-0003-3355-1848), and Rob Duffield (0000-0002-5641-1314)

Social Medica

@KyleJMBennett, @NovakSportSci, and @JobFran,

Acknowledgements

We wish to acknowledge Football Australia and its staff for their contribution to our research.

Disclosure statement

The authors report there are no competing interests to declare.

Abstract

The current study explored the association between talent pool size and relative age effects in Football Australia's talent pathway. It also compared relative age effects between male and female players. Participants were 62,717 youth football players (females: n = 15,347, age-range = 14.0 – 15.9; males: n = 47,370, age-range = 13.0 – 14.9) eligible for the National Youth Championships. We developed linear regression models to examine the association between the member federation size and the probability of a player being born earlier in the year. We also analysed selection probabilities based on birth quartile and year half across three layers. Overall, talent pool size was associated with a higher probability of selecting a player born in the first half of the year over the second. More specifically, an increase of 760 players led to a 1% higher selection probability for those born in the first six months of a chronological age group. In addition, there were more occurrences of relative age effects in the male than the female sample. Future studies should focus on the impact of the talent pool size on relative age effects at each major talent identification/selection stage of a talent pathway.

Keywords

Soccer, talent identification, player selection, talent development, selection bias

Highlights

- The study investigated the role of talent pool size and sex on relative age effects in the Australian youth football talent pathway
- Relative age effects are more prevalent in the male than female sample across (Australian) member federations.
- There was a higher probability of selecting a player born in the first half over the second half of the year for member federations with larger talent pools.

Introduction

A prominent point of discussion in football (i.e. soccer) is how to unearth players who display exceptional talent (i.e., those who can develop into elite performers in the future). However, identifying talent is difficult given the need to differentiate between a player's current performance at an early age and their potential for adult success (Baker, Schorer, & Wattie, 2018; Till & Baker, 2020). Despite the inherent challenges of assessing current performance and future potential, many national governing bodies establish talent pathways that allocate resources and coaching support towards a smaller group of youth players deemed to possess potential (Vaeyens, Güllich, Warr, & Philippaerts, 2009; Vaeyens, Lenoir, Williams, & Philippaerts, 2008; Williams, Ford, & Drust, 2020; Williams & Reilly, 2000). While talent pathways differ between national governing bodies, most entail progression from general participation in community (or grassroots) football to high-level youth and then elite senior national or international competition. Hence understanding factors within these pathways that affect the identification and development of talented players from grassroots to high-level youth competition is imperative for progressing the field.

One environmental factor associated with talent identification and selection is the relative age difference between players born close to the cut-off dates for a chronological age group (Sarmento, Anguera, Pereira, & Araújo, 2018). These chronological age differences elicit a selection or developmental bias favouring older players within an age grouping, known as relative age effects (Barnsley, Thompson, & Barnsley, 1985). In other words, players born at the start of an annual age grouping are overrepresented compared with those born near the end. There are several theoretical models (Hancock, Adler, & Cote, 2013; Wattie, Schorer, & Baker, 2015) and hypothesised explanations (Hancock, 2020) for relative age effects. For example, annual age groupings in sport often promote relative age effects by having players who differ in age by up to a year compete against each other (Cobley, Baker, Wattie, & McKenna, 2009). Further, relatively older players often possess performance advantages (possibly due to advanced physical growth and maturation) over their age-matched peers (Kelly, Côté, Jeffreys, & Tunnidge, 2021; Wattie, Cobley, & Baker, 2008). Finally, relatively older children tend to start sport earlier and accumulate more sport-specific practice and competition than younger players (Helsen, Starkes, & Van Winckel, 1998; Helsen, Van Winckel, & Williams, 2005).

Relative age effects are potentially problematic in talent pathways because relatively younger players receive fewer developmental opportunities (Ford & Williams, 2012; Helsen et al., 1998; Votteler & Höner, 2014) and are at a higher risk of dropping out (Delorme, Boiché, & Raspaud, 2010a; Lemez, Baker, Horton, Wattie, & Weir, 2014). As such, many of these relatively younger players require exceptional talent to merely 'survive' the system. While the literature extensively reports relative age effects as a confounder in the talent identification and development process in many sports and countries, the factors associated with the prevalence of relative age effects are less understood. One area requiring more attention in the research is the potential role that competition depth (i.e., talent pool size) plays in relative age effects. For example, Musch and Grondin (2001) proposed that the depth of competition for selection spots may contribute to increased prevalence of the relative age effects due to selection demands for limited places in a talent pathway. Similarly, Schorer, Cobley, Büsch, Bräutigam, and Baker (2009) reported that higher participation rates leads to greater relative age effects. Consequently, it is possible that identifying and selecting players from a large talent pool would facilitate the recruitment of more relatively older players, given their absolute number is greater than in a smaller talent pool. Additionally, increased awareness of how the size of the available selection pool may influences the occurrence of the relative age effect may help mitigate its bias in the talent identification and development process.

Another research gap is that comparatively fewer studies have focused on relative age effects in female football players (Smith, Weir, Till, Michael, & Cobley, 2018). Despite this, Delorme, Boiché, and Raspaud (2010b) identified relative age effects in French female players from under 8s to under 17s. Further, Romann & Fuchslocher (2011) observed relative age effects in Swiss talent development teams (10 – 14 years). Such a knowledge gap could lead to inefficient talent systems, sub-optimal experiences and inappropriate developmental pathways for female players (Curran, MacNamara, & Passmore, 2019). It is also pertinent to consider the respective influence of both sex and talent pool size, as some researchers have suggested that relative age effects are often less pronounced in female than male sports due to lower participation rates (Schorer et al., 2009). Though, how relative age effects appear based on sex and talent pool size across different stages of a talent pathway requires further research.

A final research gap is that many investigations focus on established football nations across Europe (Brustio et al., 2018; Del Campo, Vicedo, Villora, & Jordan, 2010; Mujika et al., 2009; Simmons & Paull, 2001; Skorski, Skorski, Faude, Hammes, & Meyer, 2016). Researchers and practitioners have less available evidence about emerging nations (Mulazimoglu, 2014; van den Honert, 2012). These nations may have smaller relative talent pools, lower absolute football participation rates, less financial and logistical resources and smaller domestic league competitions (Bennett, Vaeyens, & Fransen, 2018). Moreover, limited research has explored how relative age effects emerge across different stages of the talent pathway in emerging nations, making it challenging to optimise talent identification and development practices. Therefore, the current exploratory study offers an Australian perspective on the prevalence of relative age effects across three layers of Australia's male and female talent pathway. The study also determines the association between member federation size (i.e., general talent pool) and relative age effects in the National Premier League and National Youth Championships. Finally, the study compares relative age effects between female and male players.

Methods

Study Design

The current exploratory study adopted an observational, cross-sectional study design across four databases (three player databases based on competition level and one general population database) to facilitate three levels of analysis based on competition level. We deemed the approach necessary to reflect the hierarchical nature of talent identification and selection in Australian youth football, where coaches typically recruit players competing at higher competition levels directly from the competition level below. The study specifically compared the distributions of players' birth quartiles of the age-matched general population with all member federation registrations (Layer 1), all member federation registrations with the National Premier League Competition (Layer 2), and the National Premier League Competition with the National Youth Championships (Layer 3). We analysed these competition layers of the same age group within one national federation to determine the prevalence of the relative age effect in the talent identification process.

Australian Talent Pathway

Within Australian football, several male and female talent pathways exist via state-based (and one regional) member federations that support the national governing body (Football Australia). The overarching structure of the talent pathway across member federations and between sexes is similar, including several talent identification, selection, and development stages (Play Football, n.d.). Typically, the first stage of talent identification occurs at eight years, where players from the general talent pool (i.e., community football) trial for *skill acquisition programs* (a program developing functional game skills such as striking the ball, first touch, running with the ball, and one vs. one). The second stage of talent identification starts at 12 years when players contest spots in club academies that compete in a state-based National Premier League (highest level of weekly youth competition in most member federations). The last and arguably the most crucial stage of talent identification arises at the youth-to senior transition, which is the point where professional A-League clubs recruit players for their first or youth team (around 16 – 18 years). The talent pathway also contains two national competitions (National Youth Championships: females = 13 – 15 years, males = 13 – 14 years and National Training Centre Challenge: females = 16 years, males = 14 years) where coaches select talented players for the youth national teams. However, relative age effects within the early stages of this pathway are unknown based on the talent pool size within the respective member federations.

Participants

We were permitted to access three databases of male (born in 2005 and 2006) and female (born in 2004 and 2005) youth players registered with Football Australia and eligible to participate in the National Youth championships. Each database contained the birth dates for players at different competition levels and the nine member federations. The first database comprised players selected to compete at the 2019 National Youth Championships (females = 283 and males = 572). The second database included an age-matched sample of players registered with a club academy and competing in the 2019 National Premier League competition (females = 2,537 and males = 5,118). Two of the nine (22%) member federations did not have a National Premier League competition. One out of nine (11%) did not have National Premier League data despite having an established competition. The second database served as the National Youth Championships source population for member federations with a National Premier League competition. The third database comprised all age-matched football players registered within a member federation in Australia in that age group (i.e., the general talent pool or source population for club academies; females = 12,527 and males = 41,680). We used the third database as the National Youth Championships source population for member federations without a National Premier League competition. For the fourth and final database, we obtained an age-matched reference sample from the general population of the same years (Australian Bureau of Statistics, n.d). University of Technology Sydney's Ethics Research Committee approved the current study (ETH21-6223). Participants provided consent for Football Australia to use their data for research purposes during the registration process.

Data Analysis

All databases contained de-identified data. We classified players' birth dates according to four birth quartiles (Q1 = January, February, and March; Q2 = April, May, and June; Q3 = July, August, September; and Q4 = October, November, and December) according to Football Australia's categorisation and selection followed the standard calendar year format. We also divided players' birth dates into year halves (H1 = January to June, H2 = July to December).

Statistical Analysis

We conducted statistical analyses in RStudio (RStudio Team, 2020) and Tableau Desktop (v2021.2.2, Tableau, Seattle, Washington, USA). First, we aggregated the number of unique player identification numbers to determine the number of unique players per playing level, member federation, sex, birth quartile, and year half. Then, we determined the birth distributions as a percentage of players born within each quartile and half.

We performed comparisons at each of the three layers to determine the probability of a player being born in the first quarter or first half of the year compared with the selection pool's (the sample from where the higher layer ordinarily recruits their players) quartile or month distribution. We analysed these data using Fisher's exact test (the fisher.test function within the stats package, v3.6.2, R Core Team, 2019) from 2x2 cross-tabulations. We converted odds ratios from the model outputs to probabilities to improve interpretation. We adopted two approaches as three member federations did not have a National Premier League competition. We performed three layers of analysis for the six member federations with a National Premier League competition (layer 1 = general population vs. all member federation registrations; layer 2 = all member federation registrations vs. National Premier League; and layer 3 = National Premier League vs. National Youth Championships). We only performed two layers of analysis on remaining member federations (Layer 1 = general population vs. all member federation registrations; layer 2 modified = all member federation registrations vs. National Youth Championships).

We developed a series of linear regression models to examine the association between the member federation size and the probability of a player being born earlier in the year. In each model, the dependent variable was the probability of players being born earlier in the year at the current layer. The independent variable was the selection pool size (size of the previous layer where the higher layer ordinarily recruits their players). We performed regressions between the first and fourth quartile of the year and between the first and second half of the year, resulting in a total of four linear regressions (two layers x two conditions). Given that three member federations did not have a National Premier League competition (leaving only six data points), the male and female data were combined for the regressions, resulting in 12 data points per regression. We assessed the residual plots and QQ plots to determine model fit. We also used a Cook's distance of 4/sample-size to identify the presence of any potentially influential points. We calculated 95% confidence intervals for all model estimates. We also calculated R-squared values to determine the variance in the dependent variable the model explained.

Results

Figure 1 provides an overall description of the birth quarter and half distributions for female and male players in the Australian talent pathway across three different competition levels.

** Insert Figure 1 near here **

Selection Probabilities

Member Federations with National Premier League Data

Two member federations (females = ID 1 and 2; males = ID 1) had a lower probability of being born in Q1 than Q4 compared with the general population (Layer 1; Table 1). Relative age effects were present in four member federations (females = ID 1; males = ID 1, 2, 3 and 4) at the National Premier League level (Layer 2). Specifically, players born in Q1 had a higher selection probability than Q4 players compared with all registrations (i.e., general talent pool) for the member federation. Four member federations (males = ID 2, 3, 4, and 6) had relative age effects at the National Youth Championship level (Layer 3). One of these member federations (males = ID 6) only displayed relative age effects at the National Youth Championship level.

** Insert Table 1 near here **

Two member federations (females = ID 3; males = ID 1) had a lower probability of being born in H1 than H2 compared with the general population (Layer 1). One member federation (males = ID 5) had a higher probability of being born in H1 than H2 compared with the general population (Layer 1). Four member federations (females = ID 2; males = ID 1, 2, 3 and 4) had relative age effects at the National Premier League level (Layer 2). Three member federations (females = ID 3; males = ID 4 and 5) had relative age effects at the National Youth Championship level (Layer 3). Two of these member federations (females = ID 3; males = ID 5) only displayed relative age effects at the National Youth Championship level.

<u>**Member Federations without National Premier League Data**</u>

For member federations that did not have National Premier League competitions, one (males = ID 7) displayed a lower probability of being born in Q1 than Q4 compared with the general population (Layer 1; Table 2). Another (males = ID 9) showed relative age effects at the National Youth Championship level (Layer 2 modified).

** Insert Table 2 near here **

One member federation (females = ID 9) had a lower probability of being born in H1 than H2 compared with the general population (Layer 1). Only one member federation (males = ID 9) had relative age effects at the National Youth Championship level (Layer 2 modified).

Member Federation Size and Relative Age Effects

One member federation was substantially larger than the others and was an influential point in the dataset. We checked the model estimates with and without the influential point. These did not change substantially, and confidence intervals were slightly wider without the influential point. There was some minor skewness in the tail of the QQ plots when the influential point was present. Given that it is a true data point rather than an error and did not cause substantial changes to the model estimates, we retained this member federation in the model. The estimates for Table 3 are inclusive of all data. Regression estimates for selection pool size are small due to the scale of the data (i.e., hundreds of thousands of members). Member federation size was associated with relative age effects between the National Premier League and all member registrations for year half. An increase of 760 (95% CI: 100 – 1,410; $p = 0.028$) members within the member federation increase the probability of selecting H1 vs. H2 by 1%. Figure 2 displays scatterplots and regression lines with standard errors.

** Insert Table 4 and Figure 2 near here **

Discussion

Relative age effects remain prevalent in youth sport, though factors that affect its prevalence, such as the talent pool size in a sex-specific manner, require further research attention to address the potential problems in talent systems (Roberts, McRobert, Rudd, Enright, & Reeves, 2021). Our study provides novel insight on this topic as member federation size (i.e., talent pool size) was positively associated with the probability of selecting players born in the first half of the year compared with the second. Further, relative age effects were present at National Premier League (Layer 2) and National Youth Championship (Layer 3) levels, with these more pronounced in male than female samples. We discuss these findings further in the context of talent identification and development processes.

Our study revealed that the general participation pool showed no relative age effects, with some member federations demonstrating a lower probability of players being born in quartile one than four compared to the general Australian population. An overrepresentation towards players born right after the age cut-off was evident at higher levels of the Australian talent pathway, indicating a potential selection bias. These findings are comparable (despite our study not using a general population reference sample) to those observed in some of the most prominent global football nations. For example, Helsen et al. (2005) reported a skewed birth quartile distribution in youth national team selections (under-15, under-16, under-17, and under-18) for Belgium, England, France, Germany, Italy, the Netherlands, and Portugal. Other studies in Germany (Votteler & Höner, 2017), Spain (Mujika et al., 2009), and Italy (Brustio et al., 2018) have reported higher selection odds into talent development programs for relatively older than younger players across various youth age groups. Given these findings, the prevalence of the relative age effect across the Australian football pathway is unsurprising. Nonetheless, these relative age effects should still raise concern, especially in emerging football nations where the talent pool size is smaller, and talent identification/selection decisions have a higher risk (Bennett et al., 2018). In addition, relative age effects are associated with the dropout of relatively younger players (Delorme et al., 2010a; Helsen et al., 1998), which could further reduce emerging nations' talent pool size. Therefore, future studies should consider exploring where relative age effects arise in emerging nations 'talent pathways to offer better insight into when and how to implement future interventions (Webdale, Baker, Schorer, & Wattie, 2020).

A novel finding from the current study was the association between member federation size and relative age effects. It appears that larger talent pools increase the probability of selecting a player in the first half of the year over the second. Such a finding supports the hypothesis that competition depth may play a part in relative age effects (Musch & Grondin, 2001; Schorer et al., 2015). We note that our study assumes that all players within a talent pool have the opportunity to contest spots in talent development programs or selection into representative teams, which may not be the case. Nonetheless, the positive associations between member federation size and the probability of selecting older (first year half) vs younger players offer new insights into how the relative age effect could bias talent identification and development systems in football in any organisation with disparate feeder system sizes. It may be easier to select more relatively older players in larger member federations because more are available in the source sample (i.e., the selection pool). As a result, a singular characteristic eliciting immediate performance benefits (e.g., advanced growth and development) may bias selectors and recruiters to focus on developmental differences instead of evaluating holistic football potential. Despite this descriptive finding, the current study's results encourage further research of how talent pool size may affect the relative age effect or interventions used to mitigate it.

Another finding from our study was the fewer occurrences of relative age effects in the female sample compared with the male sample (one member federation with biases for females vs. five member federations at the highest competition level). Our findings support previous studies where relative age effects were more prevalent in male than female samples (Götze & Hoppe, 2021; Romann, Rössler, Javet, & Faude, 2018; van den Honert, 2012). However, they contrast with some studies showing relative age effects in female players (Delorme, Boiché, & Raspaud, 2010; Romann & Fuchslocher, 2011). A proposed explanation for the disparity is the earlier age of biological maturation in females compared with males (Helsen et al., 2005). Further, given that female players were older than the male players in our sample, relative age effects may have started to dissipate for females as most were likely towards the end of the adolescent growth spurt (Malina, Rogol, Cumming, Coelho-e-Silva, & Figueiredo, 2015). Another possible explanation is that relative age effects may be less noticeable in females because of lower participation (Musch & Grondin, 2001) and less competition for talent development programs (van den Honert, 2012). Due to lower participation numbers and competition, there may be a greater benefit for retaining players in programs, irrespective of whether they are relatively older or not. Together, there is a collective need for more studies to include female players to uncover the impact on the talent pathway, especially given the limited available evidence (Smith et al., 2018).

Strengths and Limitations

Although our study has addressed some gaps in the literature, there are several strengths and limitations to consider. A strength of the present study is that it included the entire population of Australian youth male and female youth football players for their respective age groups (females = 14 and 15 years and males = 13 and 14 years). The second strength of our study was our approach to data analysis, which used the selection pool as the reference point to identify relative age effects instead of the general population. Such an approach allowed us to determine the individual and sequential occurrences of relative age effects from the general talent pool (i.e., all registrations within a member federation) to the highest level of domestic competition (i.e., National Youth Championships). It also minimised the chances of overestimating the existence of relative age effects when using the general population as a reference sample for all analyses. For example, a skewed distribution towards Q1 compared to the reference sample may not be skewed if the distribution sample from which these players originate was skewed in the first place.

Despite these strengths, there were some limitations that readers should consider when interpreting the findings and implications of the research. First, the current study utilises a cross-sectional study design that captures a snapshot at one time point. As such, it does not capture the origin of relative age effects (i.e. when it becomes prevalent in the talent pathway), which might assist with the design of effective interventions. Second, the sample size of the statistical analysis for the associations between member federation size and relative age effects was small ($n = 7$). However, we used the entire population of female (13 and 15 years) and male (13 and 14 years) youth football players within Australia, so it was impossible to increase the sample size.

Conclusion

Overall, our study has added to the small number of studies on the relative age effects in female football and enabled discussions around the role of the talent pool size and relative age effects. Member federations within a talent pathway with larger talent pools were associated with a higher probability of selecting players born in the first half of the year over the second. Organisations with large talent pools should be particularly conscious of the potential for relative age effects to eventuate within pathways. Further, the female Australian youth football talent pathway showed minimal evidence for relative age effects. In comparison, there were selection biases towards relatively older players in the male sample across two competition levels. Future studies should expand on our approach and identify the role of the talent pool size at each major stage of talent identification/section.

References

Australian Bureau of Statistics. *Birth month occurrence.* https://stat.data.abs.gov.au/index.aspx?DatasetCode=BIRTHS_MONTH_OCCURRENCE, accessed 21st February 2021.

Baker, J., Schorer, J., & Wattie, N. (2018). Compromising talent: Issues in identifying and selecting talent in sport. *Quest, 70*(1), 48-63.

Barnsley, R. H., Thompson, A. H., & Barnsley, P. E. (1985). Hockey success and birthdate: The relative age effect. *Canadian Association for Health, Physical Education, and Recreation, 51*(23-28).

Bennett, K. J. M., Vaeyens, R., & Fransen, J. (2018). Creating a framework for talent identification and development in emerging football nations. *Science and Medicine in Football,* 1-7.

Brustio, P. R., Lupo, C., Ungureanu, A. N., Frati, R., Rainoldi, A., & Boccia, G. (2018). The relative age effect is larger in Italian soccer top-level youth categories and smaller in Serie A. *PLOS one, 13*(4), e0196253.

Cobley, S., Baker, J., Wattie, N., & McKenna, J. (2009). Annual age-grouping and athlete development: A meta-analytical review of relative age effects in sport. *Sports Medicine, 39*(3), 235-256.

Curran, O., MacNamara, A., & Passmore, D. (2019). What about the girls? Exploring the gender data gap in talent development. *Frontiers in Sports and Active Living, 1,* 3.

Del Campo, D. G. D., Vicedo, J. C. P., Villora, S. G., & Jordan, O. R. C. (2010). The relative age effect in youth soccer players from Spain. *Journal of sports science & medicine, 9*(2), 190.

Delorme, N., Boiché, J., & Raspaud, M. (2010a). Relative age and dropout in French male soccer. *Journal of Sports Sciences, 28,* (7), 717-722.

Delorme, N., Boiché, J., & Raspaud, M. (2010b). Relative age effect in female sport: A diachronic examination of soccer players. *Scandinavian Journal of Medicine & Science in Sports, 20,* 509-515.

Ford, P. R., & Williams, A. M. (2012). The developmental activities engaged in by elite youth soccer players who progressed to professional status compared to those who did not. *Psychology of Sport and Exercise, 13*(3), 349-352.

Götze, M., & Hoppe, M. W. (2021). Relative age effect in elite german soccer: influence of gender and competition level. *Frontiers in Psychology, 11*, 3725.

Hancock, D.J. (2020). Antecedents and explanations of relative age effects. In J.C. Dixon, S. Horton, L. Chittle & J. Baker, *Relative age effects: International Perspectives* (pp. 23-32). Routledge.

Hancock, D.J., Adler, A.L., & Côté, J. (2013). A proposed theoretical model to explain relative age effects in sport. *European Journal of Sport Sciences, 13*(6), 630-637.

Helsen, W. F., Starkes, J. L., & Van Winckel, J. (1998). The influence of relative age on success and dropout in male soccer players. *American Journal of Human Biology, 10*(6), 791-798.

Helsen, W. F., Van Winckel, J., & Williams, A. M. (2005). The relative age effect in youth soccer across Europe. *Journal of Sports Sciences, 23*(6), 629-636.

Kelly, A. L., Côté, J., Jeffreys, M., & Tunnidge, J. (2021). *Birth advantages and relative age effects in sport: Exploring organizational structures and creating appropriate settings*: Routledge.

Lemez, S., Baker, J., Horton, S., Wattie, N., & Weir, P. (2014). Examining the relationship between relative age, competition level, and dropout rates in male youth ice-hockey players. *Scandinavian Journal of Medicine & Science in Sports, 24,* 935-942.

Malina, R. M., Rogol, A. D., Cumming, S. P., Coelho-e-Silva, M. J., & Figueiredo, A. (2015). Biological maturation of youth athletes: Assessment and implications. *British Journal of Sports Medicine, 49*(13), 852-859.

Mujika, I., Vaeyens, R., Matthys, S. P. J., Santistenban, J., Goiriena, J., & Philippaerts, R. (2009). The relative age effect in a professional football club setting. *Journal of Sports Sciences, 27*(11), 1153-1158.

Mulazimoglu, O. (2014). The relative age effect (RAE) in youth and professional soccer players in Turkey. *The Anthropologist, 18*(2), 391-398.

Musch, J., & Grondin, S. (2001). Unequal competition as an impediment to personal development: A review of the relative age effect in sport. *Developmental review, 21*(2), 147-167.

Play Football (n.d.). *A guide to the development pathways available to male and female football players in Australia.* https://www.playfootball.com.au/pathways.

Roberts, S. J., McRobert, A. P., Rudd, J., Enright, K., & Reeves, M. J. (2021). Research in Another un-Examined (RAE) context. A chronology of 35 years of relative age effect research in soccer: is it time to move on? *Science and Medicine in Football, 5*(4), 1-9.

Romann, M., & Fuchslocher, J. (2011). Influence of the selection level, age and playing position on relative age effects in Swiss women's soccer. *Talent Development & Excellence, 3*(2), 239-247.

Romann, M., Rössler, R., Javet, M., & Faude, O. (2018). Relative age effects in Swiss talent development–a nationwide analysis of all sports. *Journal of Sports Sciences, 36*(17), 2025-2031.

Sarmento, H., Anguera, M. T., Pereira, A., & Araújo, D. (2018). Talent identification and development in male football: A systematic review. *Sports Medicine, 48*(4), 907-931.

Schorer, J., Cobley, S., Büsch, D., Bräutigam, H., & Baker, J. (2009). Influences of competition level, gender, player nationality, career stage and playing position on relative age effects. *Scandinavian Journal of Medicine & Science in Sports, 19*(5), 720-730.

Schorer, J., Cobley, S., Bräutigam, H., Loffing, F., Hütter, S., Büsch, D., . . . Baker, J. (2015). Developmental contexts, depth of competition and relative age effects in sport: a database analysis and a quasi-experiment. *Psychological Test and Assessment Modeling, 57*(1), 126.

Simmons, C., & Paull, G. C. (2001). Season-of-birth bias in association football. *Journal of Sports Sciences, 19*(9), 677-686.

Skorski, S., Skorski, S., Faude, O., Hammes, D., & Meyer, T. (2016). The relative age effect in elite German youth soccer: Implications for a successful career. *International Journal of Sports Physiology and Performance, 11*(3), 370-376.

Smith, K. L., Weir, P. L., Till, K., Michael, R., & Cobley, S. (2018). Relative age effects across and within female sport contexts: A systematic review and meta-analysis. *Sports Medicine, 48*, 1451-1478.

Till, K., & Baker, J. (2020). Challenges and [possible] solutions to optimizing talent identification and development in sport. *Frontiers in psychology, 11*, 664.

Vaeyens, R., Güllich, A., Warr, C. R., & Philippaerts, R. (2009). Talent identification and promotion programmes of Olympic athletes. *Journal of Sports Sciences, 27*(13), 1367-1380.

Vaeyens, R., Lenoir, M., Williams, A. M., & Philippaerts, R. M. (2008). Talent identification and development programmes in sport: Current models and future directions. *Sports Medicine, 38*(9), 703-714.

van den Honert, R. (2012). Evidence of the relative age effect in football in Australia. *Journal of Sports Sciences, 30*(13), 1365-1374.

Votteler, A., & Höner, O. (2014). The relative age effect in the German Football TID Programme: Biases in motor performance diagnostics and effects on single motor abilities and skills in groups of selected players. *European Journal of Sport Sciences, 14*(5), 433-442.

Votteler, A., & Höner, O. (2017). Cross-sectional and longitudinal analyses of the relative age effect in German youth football. *German Journal of Exercise and Sport Research, 47*(3), 194-204.

Wattie, N., Cobley, S., & Baker, J. (2008). Towards a unified understanding of relative age effects. *Journal of Sports Sciences, 26*(13), 1403-1409.

Wattie, N., Schorer, H., & Baker, J. (2015). The relative age effect in sport: A developmental systems model. *Sports Medicine, 45*, 83-94.

Webdale, K., Baker, J., Schorer, J., & Wattie, N. (2020). Solving sport's 'relative age' problem: A systematic review of proposed solutions. *International Review of Sport and Exercise Psychology, 13*(1), 187-204.

Williams, A. M., Ford, P. R., & Drust, B. (2020). Talent identification and development in soccer since the millennium. *Journal of Sports Sciences, 38*(11-12), 1199-1210.

Williams, A. M., & Reilly, T. (2000). Talent identification and development in soccer. *Journal of Sports Sciences*, *18*(9), 657-667.

Tables

Table 1. The probability (95% CI) of male and female youth players being born in Q1 vs. Q4 and H1 vs. H2 relative to the selection pool for member federations with National Premier League data

Playing Level	Member Federation	Layer 1		Layer 2		Layer 3	
		Probability	**p-value**	**Probability**	**p-value**	**Probability**	**p-value**
Q1 vs. Q4							
Females	ID 1 (n = 5,604)	48% (46-50%)	**0.037**	53% (47-59%)	0.338	64% (43-82%)	0.174
	ID 2 (n = 2,062)	47% (44-50%)	**0.035**	68% (57-79%)	**0.001**	39% (13-74%)	0.512
	ID 3 (n = 1,675)	49% (46-53%)	0.754	56% (46-66%)	0.277	73% (39-94%)	0.236
	ID 4 (n = 686)	50% (45-56%)	> 0.999	47% (38-56%)	0.537	49% (20-77%)	> 0.999
	ID 5 (n = 541)	47% (41-53%)	0.343	49% (37-62%)	> 0.999	42% (15-73%)	0.767
	ID 6 (n = 498)	51% (45-58%)	0.655	53% (39-67%)	0.680	39% (11-77%)	0.725
Males	ID 1 (n = 15,803)	48% (47-49%)	**< 0.001**	64% (60-68%)	**< 0.001**	56% (60-68%)	0.484
	ID 2 (n = 7,359)	50% (49-52%)	0.854	67% (62-71%)	**< 0.001**	76% (54-91%)	**0.016**
	ID 3 (n = 6,657)	50% (48-52%)	0.775	64% (58-69%)	**< 0.001**	78% (60-90%)	**0.001**
	ID 4 (n = 3,788)	51% (49-53%)	0.404	60% (54-67%)	**0.002**	77% (52-93%)	**0.029**
	ID 5 (n = 2,016)	53% (50-56%)	0.066	52% (47-57%)	0.383	57% (37-76%)	0.586
	ID 6 (n = 1,317)	48% (44-52%)	0.389	56% (47-65%)	0.167	91% (61-100%)	**0.005**
H1 vs. H2							
Females	ID 1 (n = 5,604)	49% (48-51%)	0.391	51% (46-55%)	0.797	58% (43-72%)	0.251
	ID 2 (n = 2,062)	49% (46-51%)	0.193	61% (53-68%)	**0.007**	55% (33-76%)	0.834
	ID 3 (n = 1,675)	47% (45-50%)	**0.031**	52% (45-59%)	0.526	71% (50-85%)	**0.042**
	ID 4 (n = 686)	50% (47-54%)	0.819	47% (41-53%)	0.354	57% (36-75%)	0.560
	ID 5 (n = 541)	49% (45-53%)	0.667	48% (40-56%)	0.667	50% (28-72%)	> 0.999
	ID 6 (n = 498)	50% (45-54%)	0.858	52% (42-61%)	0.709	41% (22-62%)	0.419
Males	ID 1 (n = 15,803)	49% (48-50%)	**0.014**	61% (58-64%)	**< 0.001**	57% (46-68%)	0.206
	ID 2 (n = 7,359)	51% (50-52%)	0.232	61% (57-64%)	**< 0.001**	54% (41-67%)	0.539
	ID 3 (n = 6,657)	51% (49-52%)	0.264	59% (55-62%)	**< 0.001**	63% (52-74%)	0.025
	ID 4 (n = 3,788)	51% (49-53%)	0.216	59% (55-64%)	**< 0.001**	72% (58-84%)	**0.002**
	ID 5 (n = 2,016)	53% (51-55%)	**0.012**	50% (47-54%)	0.827	65% (52-77%)	**0.023**
	ID 6 (n = 1,317)	50% (47-52%)	0.869	54% (48-60%)	0.171	72% (51-87%)	0.025

Notes: 50% is the expected distribution based on the selection pool. Layer 1 = all member federation registrations vs. an age-matched general population, Layer 2 = National Premier League vs. all member federation registrations, and Layer 3 = National Youth Championships vs. National Premier League.

Table 2. The probability (95% CI) of male and female youth players being born in Q1 vs. Q4 and H1 vs. H2 relative to the selection pool for member federations without National Premier League data

Playing Level	Member Federation	Layer 1		Layer 2	
		Probability	**p-value**	**Probability**	**p-value**
Q1 vs. Q4					
Females	ID 7 (n =70)	51% (34-67%)	> 0.999	79% (25-99%)	0.356
	ID 8 (n = 207)	55% (45-65%)	0.316	55% (21-86%)	> 0.999
	ID 9 (n = 1,184)	47% (43-51%)	0.174	55% (31-78%)	0.816
Males	ID 7 (n = 227)	39% (30-48%)	**0.018**	64% (36-85%)	0.299
	ID 8 (n = 801)	47% (42-52%)	0.289	57% (33-79%)	0.683
	ID 9 (n = 3,812)	49% (46-51%)	0.212	85% (66-96%)	**< 0.001**
H1 vs. H2					
Females	ID 7 (n =70)	49% (37-62%)	> 0.999	74% (38-95%)	0.197
	ID 8 (n = 207)	52% (45-59%)	0.488	53% (33-73%)	0.843
	ID 9 (n = 1,184)	46% (44-49%)	**0.016**	49% (35-63%)	0.892
Males	ID 7 (n = 227)	46% (39-52%)	0.207	55% (35-73%)	0.705
	ID 8 (n = 801)	50% (46-53%)	0.832	53% (35-70%)	0.863
	ID 9 (n = 3,812)	49% (47-51%)	0.230	72% (59-82%)	**< 0.001**

Note: 50% is the expected distribution based on the selection pool. Layer 1 = all member federation registrations vs. an age-matched general population, Layer 2 = National Youth Championships vs. all member federation registrations.

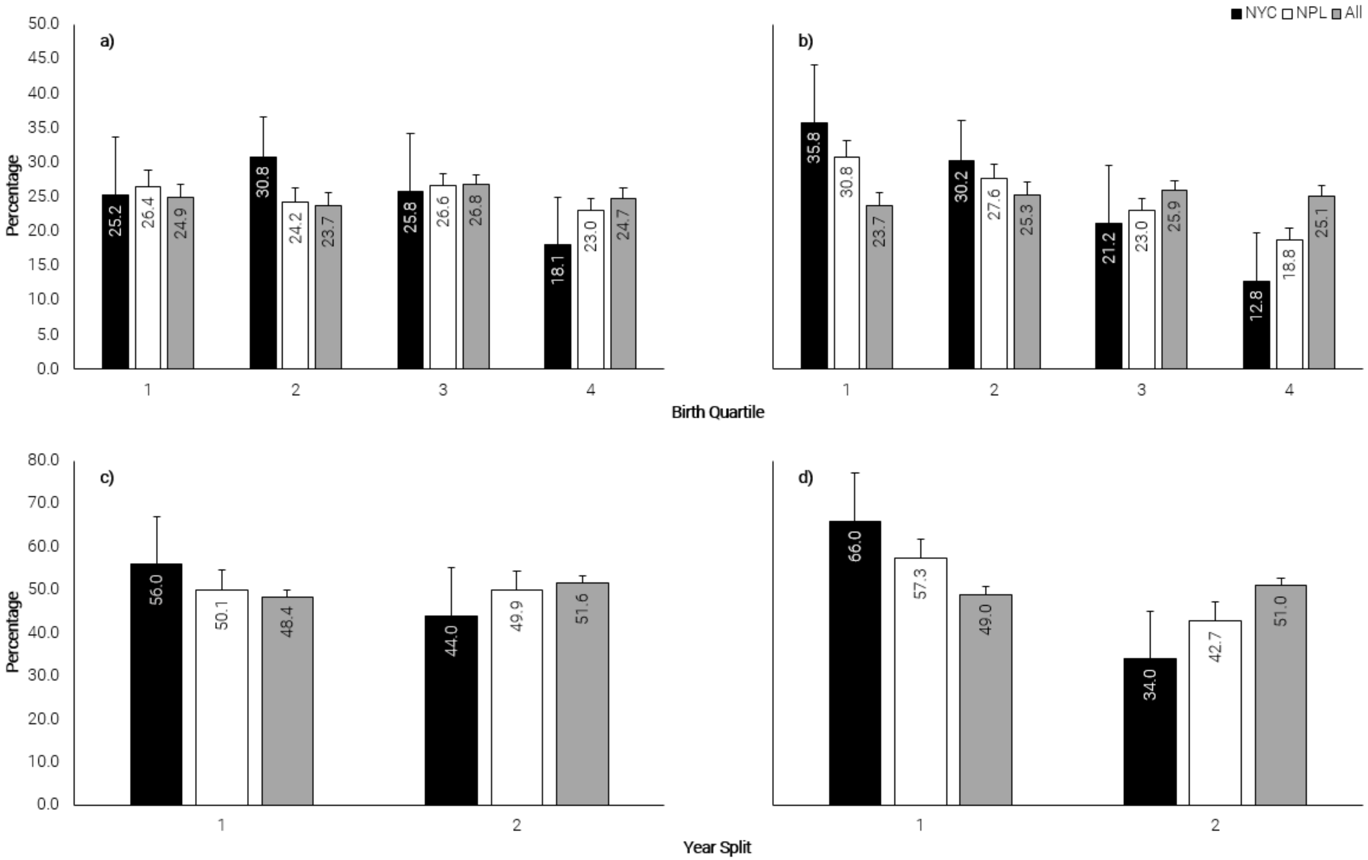
Table 3. Linear regression models to determine associations between selection pool size and probability of selection in earlier halves and quarters of the year

Model	Model Parameter	Estimate (95% CI)	p	R^2
H1 vs H2 NPL from ALL	Intercept	51.6 (47.7-55.4)	<0.001	0.40
	Selection Pool Size	0.00076 (0.0001-0.00141)	0.028	
Q1 vs Q4 NPL from ALL	Intercept	53.8 (48.4-59.2)	<0.001	0.32
	Selection Pool Size	0.00091 (-0.00002-0.00183)	0.054	
H1 vs H2 NYC from NPL	Intercept	51.8 (45.9-57.7)	<0.001	0.14
	Selection Pool Size	0.00467 (-0.00358-0.01293)	0.236	
Q1 vs Q4 NYC from NPL	Intercept	55.5 (35.4-75.5)	<0.001	0.07
	Selection Pool Size	0.01066 (-0.01743-0.03875)	0.418	

Notes: NYC = National Youth Championship, NPL = National Premier League, All = all players in member federation

Figures

Figure 1. Birth quarter and half distributions for female (born 2004-2005) and male (born 2005-2006) players registered with football Australia across three different competition levels (National Youth Championships [NYC], National Premier League [NPL] and an age-matched reference sample from the general participation pool.



Notes: a) Female birth quartile distribution, b) Male birth quartile distribution, c) Female year half, and d) Male year half. NYC = National Youth Championship (females: n = 283, males: n = 572), NPL = National Premier League (females: n = 2,537, males: 5,118), All = all players in member federation (females: n = 12,527, males: n = 41,680).

Figure 2. Associations between the probability of being selected in H1 vs. H2 or Q1 vs. Q4 and the selection pool size.