UNIVERSIDADE D
COIMBRA

João Ricardo Horta Rocha da Palma

ASSOCIAÇÃO DO POLIMORFISMO rs35767 DO GENE *IGF1* COM A PERFORMANCE DESPORTIVA: UMA META-ANÁLISE

Dissertação no âmbito do Mestrado em Biocinética, orientada pelo Professor Doutor Amândio Cupido Manuel dos Santos e pela Professora Doutora Henriqueta Alexandra Mendes Breda Lobo Coimbra Silva e apresentada na Faculdade de Ciências do Desporto e Educação Física

Setembro de 2023

The whole of science is nothing more than a refinement of everyday thinking.

Albert Einstein

Acknowledgements

First of all, I would like to thank everyone who in some way or another has influenced the writing of this thesis.

My special thanks goes to my supervisor, Professor Amândio Manuel Cupido dos Santos, for taking on this project in its earliest and most volatile phase, providing me the guidance necessery for me to find the path that would allow me to finish this work.

I could not have taken on this journey without the invaluable contribuitons of my supervisor Professor Henriqueta Alexandra Mendes Breda Lobo Coimbra Silva and of Professor João Fernando Pereira Mendes, who have provided with all of the knowledge and most importantly, their time, patience and unrelenting effort day in and day out to make sure that this thesis was completed in the timetable avaliable. I would also like to express my utmost gratitude to Professor Bárbara Cecília Bessa dos Santos Oliveiros Paiva, for her contributions came in many, sometimes late, hours of need.

My many thanks must also go to my partner, Francisca, for all of the hours that she spent listening, helping and simply for making me push harder to be a better person every day.

And lastly, but surely not least important, my biggest gratitude goes to my mother, Custódia, who through thick and thin was always there for me, never judging and always supportive, either with a word of advice or a simple gesture of care. For all the examples of resilience, effort and keeping a calm head in hours of more stress, I thank you.

Resumo

Introdução: A performance atlética é um fenótipo multifatorial dependente da interação de múltiplas variantes genéticas e não genéticas. Embora mais de 200 variantes genéticas tenham já sido associadas à variabilidade interindividual da performance atlética, muitos dos estudos realizados, particularmente os de tipo "gene candidato", têm uma amostra de pequena dimensão e são inconclusivos, condicionando a utilização do perfil génico como uma ferramenta para a seleção de atletas e refinamento de programas de treino individualizados. O IGF1 é um fator de crescimento que está envolvido no desenvolvimento e na função de massa muscular esquelética, sendo considerado um agente dopante no mundo do desporto. O polimorfismo de nucleótido único (SNP) rs35767, localizado na região reguladora do gene *IGF1*, influência os níveis de síntese da respetiva proteína, e tem vindo a ser associado a fenótipos atléticos. O objetivo deste trabalho for realizar uma meta-análise para avaliar a associação entre o polimorfismo rs35767 do gene *IGF1* e a performance em desportos de endurance e potência/força.

Metodologia: Foi recolhida literatura das bases de dados PubMed, Web of Science, Cochrane, Scopus e Embase, até ao mês de agosto de 2023. O "odds ratio" (OR) e o Intervalo de Confiança de 95% (CI) foram obtidos através da utilização do modelo de efeitos fixos e permitiram as comparações entre cada grupo de atletas e controlos e entre os dois grupos de atletas. Diferenças estatisticamente significativas foram consideradas a partir de $p<0.05$.

Resultados: Três estudos, com um total de 354 atletas (de nível elite ou nacional) e 234 controlos foram incluídos nesta meta-análise. O equilíbrio de Hardy-Weinberg for verificado em cada estudo e para o total dos controlos ($p>0.05$). A frequência do alelo minor foi 0.12 nos controlos, 0.21 em atletas de endurance, 0.19 em atletas de potência/força. Diferenças estatisticamente significativas foram encontradas entre os grupos de endurance e controlos ($p= 0.005$; OR = 1.91; CI = 1.21-3.01) e entre atletas de potência/força e controlos ($p = 0.01$; OR= 1.57; CI = 1.11-2.24). Não foram encontradas diferenças estatisticamente significativas entre os grupos de potência/força e endurance.

Conclusões: Esta meta-análise apoia a hipótese de o alelo minor do polimorfismo rs35767 do gene *IGF1* favorecer a performance atlética em desportos de endurance e potência/força.

Palavras-Chave: rs35767, IGF-1, polimorfismo, atletas, desporto de potência/força; desporto de endurance, fenótipo multifatorial

Abstract

Background: Sport performance is a multifactorial phenotype dependent on the interaction of multiple genetic and non-genetic factors. More than 200 polymorphisms have been associated with athletic performance, but most of these studies have a small sample size, and are inconclusive, hampering the use of genetic profiles as an effective tool for athlete selection and refinement of personalized training programs. IGF1 is a growth factor involved in skeletal muscle development and function, and is considered a sport doping agent. The single nucleotide polymorphism (SNP) rs35767, located in the regulatory region of *IGF1* gene, influences its expression and has been associated with sports-related phenotypes. Our aim was to perform a meta-analysis to evaluate the association between the rs35767 polymorphism of the *IGF1* gene and athletic performance in power and endurance sports.

Methods: Literature has been retrieved from PubMed, Web of Science, Cochrane, Scopus, and Embase databases until August 2023. The odds ratio (OR) and 95% confidence interval (CI) were obtained using the fixed-effects model for comparisons between power sports athletes and controls, endurance sports athletes and controls, and between the two groups of athletes. Statistical significance was defined for $p<0.05$.

Results: Three studies, with a total of 354 athletes (top or national level) and 234 controls were included in this meta-analysis. Hardy Weinberg equilibrium was verified in each study and for the total of controls ($p>0.05$). The frequency of minor allele was 0.12 in controls, 0.21 in endurance athletes, and 0.19 in power sport athletes. Statistically significant differences were highlighted for the frequency of the minor allele between endurance athletes and controls ($p= 0.005$; OR = 1.91; CI = 1.21-3.01) and between power sport athletes and controls ($p = 0.01$; OR= 1.57; CI = 1.11-2.24). No statistically significant difference was found between the power and endurance groups

Conclusions: This metanalysis supports the role of the minor allele of the rs35767 polymorphism of the *IGF1* gene as favoring an athlete's performance in endurance and power sport.

Keywords: rs35767, IGF-1, Polymorphism, Athletes; power sport; endurance; multifactorial phenotype.

Table of Contents

Abbreviations

ACE - Angiotensin-converting enzyme

ACTN3 - α- actinin 3

GWAS - Genome-wide association studies

IGF1 - Insuline like Growth factor 1

IGF2 - Insuline like Growth factor 2

NOS - Endothelial nitric oxide synthase 3

SNP's - Single Nucleotide Polymorphism

MYBPC3 - Myosin-binding protein *C*

CDKN1A - cyclin dependent kinase inhibitor 1A

SIRT1 - NAD-dependent deacetylase sirtuin-1

PPARG - peroxisome proliferator activated receptor gamma

DNA - Deoxyribonucleic Acid

IGF1R - Insulin like growth factor 1 receptor

IGF1BP - Insulin-like growth factor-binding protein

List of Tables

List of Figures

Introduction

In the current sport world, optimized protocols from diverse areas such as nutrition, sports medicine, personalized and specific physical and psychological training, and monitoring strategies are applied in what became a new science of "Talent development" (Vaeyens, R et al. 2009). Identifying juveniles with potential to became high performance athletes naturally became a major issue in this highly competitive field.

Athletic success is a phenotype with a complex and dynamic architecture, including many interdependent factors such as biological, and psychological to training. Among biological factors, individual genetic profile is supposed to have a major influence in determining the potential of each athlete's performance and response to training (Semenova, E.A et al. 2023; Sellami 2022; Varillas-Delgado D. et al. 2022). Association studies have identified genetic variants, mostly single nucleotide polymorphisms (SNPs), directly or indirectly associated with sports-related phenotypes. Candidate gene approaches, that explore genes selected by their hypothetical role in the evaluated trait, were the first to identify genetic variants associated with elite athlete status (Bouchard C. et al. 1988) Later, advances in genomic technology allowed broader, hypothesis-free, genomic variants' search, using array-based SNP identification (Genome-Wide Association Studies or GWAs), genome sequencing and exome sequencing (Ahmetov, I. et al. 2015; Al-Khelaifi, F et al. 2020; Bulgay, C. et al. 2023, Semenova, E.A et al. 2023). Globally, many genetic variants were described in genes encoding proteins with different functions such as growth factors (*IGF1* and *IGF2*), blood pressure regulation and electrolyte balance (*ACE*), heart muscle cell's function (*MYBPC3*), cell cycle regulation (*CDKN1A*), control of gene expression, metabolism and aging (*SIRT1*), transcription factor regulator of adipocyte differentiation (*PPARG*), skeletal muscle function (*ACTN3*) or nitric oxide synthesis (*NOS3*), among others (review in Ekaterina A. Semenova 2023 Update). According to twin studies estimate, the heritability of athlete performance ranges from 0.27 to 0.78. (Beunen & Thomis, 1999; De Moor et al., 2007; Eriksson et al., 2006; Maia et al., 2002; Pittaluga et al., 2004). It has now been recognized that hundreds of variants may be involved and that they are characterized, in their majority, by a population frequency higher than 0.05, being classified as DNA polymorphisms, and low penetrance, explaining why they are

so difficult to identify (Varillas-Delgado D. 2022) and the requirement for large population samples.

Gene-gene, epigenetic and other and gene-environment interactions, which are not static, create significant additional challenges (Kashimoto 2016; Sellami 2021; Varillas-Delgado 2022).

Insulin-like growth factor 1 (IGF1), encoded by *IGF1* gene, is similar to insulin in structure and function and is the primary mediator of the effects of growth hormone. With a peak expression in pubertal growth spurt, it continues to have an anabolic effect in adults. The liver is the main source of plasma IGF1, though it also expressed by most tissues where it acts in a paracrine and autocrine way. Binding of IGF1 to its transmembrane receptor IGF1R, results in the activation of multiple signaling pathways including the PI3K/Akt pathway and the Raf/MEK/ERK cascade that inhibit apoptosis and promote cellular growth and survival (Gusscott 2016). The role of IGF1 in skeletal muscle and bone development, growth, and strength is well documented and supported by animal and human studies (Bikle 2015; Gonzalez 2003; Ostrovsky O 2010; Velloso 2008; Wang 2011; Yakar 2018). Human studies have associated IGF1 expression with sports-related phenotypes such as isometric exercise in healthy human (Creig 2006), physical activity fitness in adolescent females (Eliakim 1996;), short/long-term resistance-training programs [Roelen, C.A. et al. 1997, Rubin, M.R. et al. 2005], high-volume training (Koziris, L.P. et al. 1999,Marx J.O. et al. 2001), triathlon training [Maimoun L. et al. 2004) and body composition and physical performance in female Olympic athletes (Eklund E, 2021), among others. Moreover, IGF1 is considered to be a doping agent in sport (Ernst, 2013; García-Arnés 2023).

IGF1 expression is regulated at different levels, from transcription to transcript processing, with the occurrence of different splice-variants, protein modifications and complexing to IGF1 binding proteins (IGF1BP) (review in Bailes 2021). The SNP rs35767 (previously described as -C1245T) is located in the regulatory region of the *IGF1* gene. The allele G (or C depending on the DNA strand of reference) is the more frequent variant worldwide, and allele A (or T depending on the DNA strand of reference) is the more frequent alternative allele (SNP data base; https://www.ncbi.nlm.nih.gov/snp/rs35767). There is a growing body of evidence suggesting that the minor allele is associated with increased circulating levels of IGF1 (Palles C 2008; Canzian 2006; Ben-Zaken S. et al. 2015) and with augmented muscle

mass. (Ben-Zaken S. et al. 2013;). Kostek and collaborators found that elderly Caucasian women homozygotes for the major allele show an increase in body fat mass of 2%, when compared to heterozygotes (Kostek M.C. et al. 2010).

On the other hand, there are studies showing an association between the minor allele and better performance in different sports, from power to endurance and others (Ben-Zaken S. et al. 2013, 2014, 2022, Karpowicz K. et al. 2018, Ben-Zaken S. et al. 2017, Batavani M.R. et al. 2018, Moreland,E. et al. 2022). Generally, the statistical power of these studies is hampered by the small size of sample population.

The aim of this study is to perform a systematic review and meta-analysis to evaluate if rs35767 polymorphism is associated with athletic performance in endurance and power sports.

Methods

Protocol and registration

This study was designed according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement. A registration with the ID 463864 was made in the PROSPERO database was made at inception.

Literature search

This study was performed according to the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines. The bibliographical search was conducted in PubMed, Web of Science, Cochrane, Scopus and Embase, up to 15 August 2023 using the following keywords and boolean operators: "IGF1 (**OR** IGF-1 **OR** Insulin-like growth factor 1) **AND** rs35767 (**OR** rs60611542 **OR** rs17884459 **OR** -C1245T) **AND** Athletes. All papers retrieved were exported to Endnote Citation Management Software, and duplicates were erased. References of selected papers were checked for other relevant publications. Figure 1 shows the search strategy applied in this systematic review.

Inclusion and exclusion criteria

The inclusion criteria were: be written in English, French, Portuguese, or Spanish; be human case-control studies; include analysis of rs35767; were conducted on elite or national level athletes; include athletes of endurance and power sports (individual sports); had a control sample of healthy non-athlete individuals; include more than 20 athletes and more than 20 controls, referred submission to ethical approval committee.

The exclusion criteria were: animal or *in vitro* studies; studies with individuals younger than age 16; do not include a control population of healthy non-athlete individuals; were reviews or metanalysis, evaluated clinical conditions and not athletic phenotypes; do not include the rs35767 polymorphism; whose data was incomplete or did not allowed further analysis; Hardy-Weinberg equilibrium was not evaluated.

Studies selection and data extraction

Two authors (PJ, SHC) systematically screened the titles, abstracts and full text (by this order) of retrieved publications, applying the inclusion and exclusion criteria. Disagreement was resolved by consensus and referral to a third author (MJF) if necessary.

Classification of sports as being endurance or power sport, was according to (Fink, H. H., & Mikesky, A. E. 2017, p.362-363, Bompa & Buzzichelli, 2021, p.40-43). There was no blinding to study author, institution, or journal throughout. Data collected included: authors, date of publication and title, population origin of cases and controls, type of sport, elite or national level athlete, age and number of each genotype in cases and controls. Three authors (JFM, PJ, SHC) checked the data extraction.

Qualitative assessment

Two authors independently assessed the risk of bias applying the Newcastle-Ottawa quality assessment scale (NOS).

Statistical Analysis

After the systematic review was completed, a meta-analysis was performed to evaluate the association between the rs35767 with athletic performance in endurance and power sports. Heterogeneity between the studies was calculated with the I^2 statistic (Higgins, J. et al. 2003). The allelic frequency of the minor allele was compared between cases (athletes) and controls (allelic model). The odds ratio (OR) and its 95% confidence interval (CI) were used to measure the strength of the association. A random-effects model, using fixed-effects model, was used to calculate pooled ORs. The significance of pooled ORs was determined with Z tests ($p < 0.05$ was considered statistically significant). Compliance with the Hardy-Weinberg equilibrium (HWE) in controls was tested via a Chi-square test with statistical significance set to $p<0.05$. Statistical analysis was performed with RevMan 5.4.1

Results

After the removal of all duplicates, 245 articles were retrieved. The selection process is described in Figure 1.

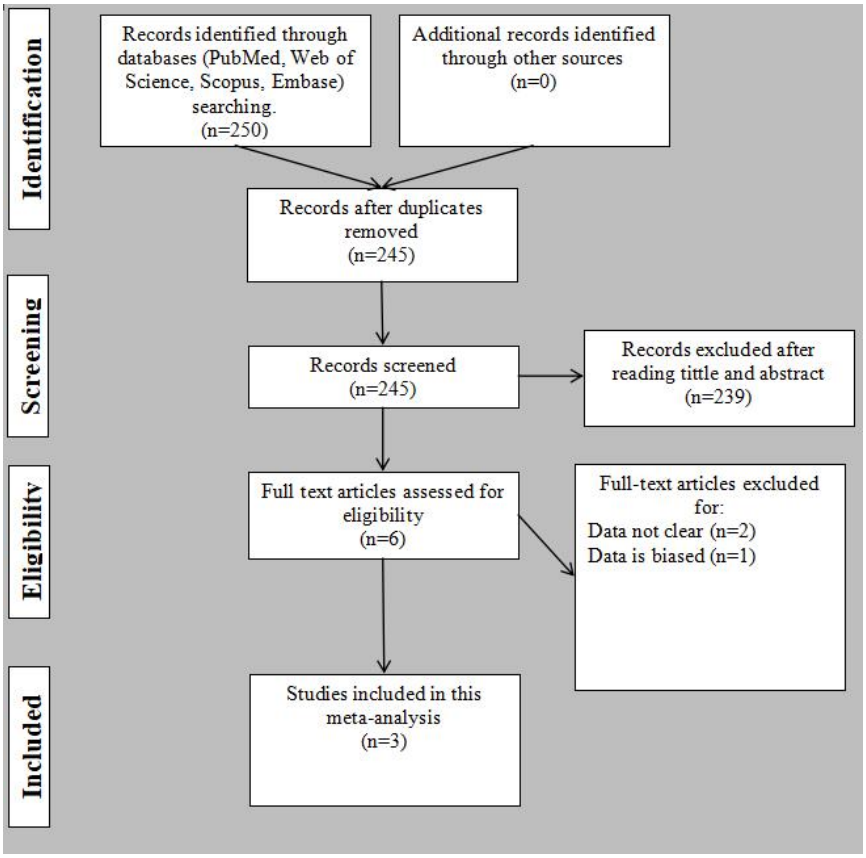

Figure 1 - Flow chart for the Systematic Review

The title and abstract of the 245 unique publications were evaluated to verify compliance with inclusion and exclusion criteria, resulting in the removal of 239 publications. Six articles were submitted for a full text evaluation, and three were removed for specific reasons: data insufficient to evaluate genotypes and alleles frequency of rs35767 (Ben-Zaken S. et al. 2017, Batavani M.R. et al. 2018) and only data corresponding to results with statistically significant difference were described (Moreland,E. et al. 2022). For the three studies included in the meta-analysis, a risk of bias analysis was conducted using the Newcastle-Ottawa Scale (NOS) (Table 1).

Table 1- Evaluation of risk of bias applying the Newcastle-Ottawa scale

References	Selection	Comparability	Outcome	Score*
Ben-Zaken S. et al. 2013	****	*	**	7
Ben-Zaken S. et al. 2014	****	*	**	7
Ben-Zaken S. et al. 2022	****	*	**	7

*Newcastle-Ottawa's scale publications are classified as: high quality (7-9), high risk of bias (4-6) and very high risk of bias (0-3).

The three selected studies comprised a total of 354 elite or national level athletes, (116 endurance athletes and 238 power sport athletes) and 234 controls. Gender distribution was of 18.1 % for women and 81.8% for men in the athlete group and of 28.4% for women and 76.1% for the men in the control group. Of the athletes described in these studies, only those practicing a sport classified as endurance (1500m running; half Ironman; 400-1500m swimming) or as power sport (decathlon, weightlifting, 100-200m runs, long jump, 50-100m swimming) were analyzed (Table 2).

Table 2 - Number of individuals per sport

References	Sport	Number of athletes
Ben-Zaken S. et al., 2013	Endurance Type[1]	78
	Power Type[2]	87
Ben-Zaken S. et al., 2022	Sprint and Long Jump	40
	MDR*	38**
	Weightlifting	44
	Decathlon	27
Ben-Zaken S. et al., 2014	Short Distance Swimming	42
	Long Distance Swimming	38

1- 1500m to 113.13 km (half-ironman); 2- 100-200m runs and Long Jump; *- Middle Distance Running
**- Data excluded from the meta-analysis

The group of athletes of 800-1500m distance running (MDR) referred by Ben-Zaken and collaborators (2022) was not analyzed as its classification as endurance or power sport is controversial. It was not possible to analyze elite and national level athletes as this information was not available in all papers. Major retrieved data is described in Table 3.

Table 3 - Global data of the three studies

		Cases						Controls				
			Genotypes				Age	Genotypes				Age
Ref	**Origin of population**	**Sport type**	**CC**	**CT**	**TT**	**N**	average±sd; (Years)	**CC**	**CT**	**TT**	**N**	average±sd (Years)
1	Israeli	Power sport*	59	24	4	87	35.9±12.2	119	40	0	159	26±3.0
		Endurance	49	25	4	78						
2	Israeli	Power sport*	29	13	0	42	23.1±7.6	119	40	0	159	26.4±5.8
		Endurance	23	13	2	38	25.3±9.3					
3	Israeli and Estonian	Power sport – Sprint and Long Jump	23	13	4	40	34.1±15.5	54	20	1	75	28±8.7
		Power sport – MDR	21	17	0	38	29.9±11.2					
		Power sport – Weightlifiting	28	14	2	44	33.9±11.1					
		Power sport – Decathlon	15	9	3	27	38.1±14.1					

1.Ben-Zaken S. *et al*. 2013; 2. Ben-Zaken S. *et al*.2014; 3- Ben-Zaken S. *et al*.2022; N – total number of individuals; sd- standard deviation

The value for heterogeneity across the studies was adequate (I^2 =0). The frequency of minor allele was 0.12 in controls, 0.21 in endurance athletes and 0.19 in power sport athletes. As in Ben-Zaken *et al*, 2013 and in Ben-Zaken *et a*l 2014, the controls were the same (as described in table 3), to allow a correct statistical analysis removing the duplication of controls, the two papers were analyzed as one, merging athletes of the same type of sport (power or endurance) as visualized in Forest charts (Figures 2 and 3).

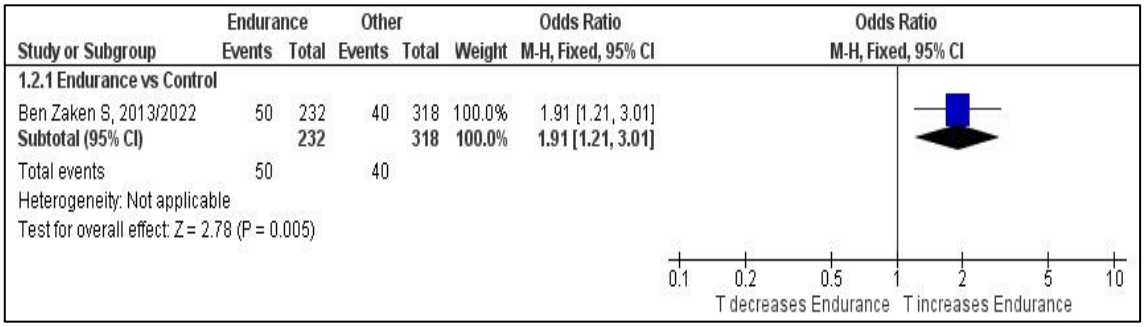

Figure 2 – Forest chart comparing endurance athletes and controls. The results correspond to the merge of Ben-Kazen et al 2013 and Ben-Zaken et al 2014.

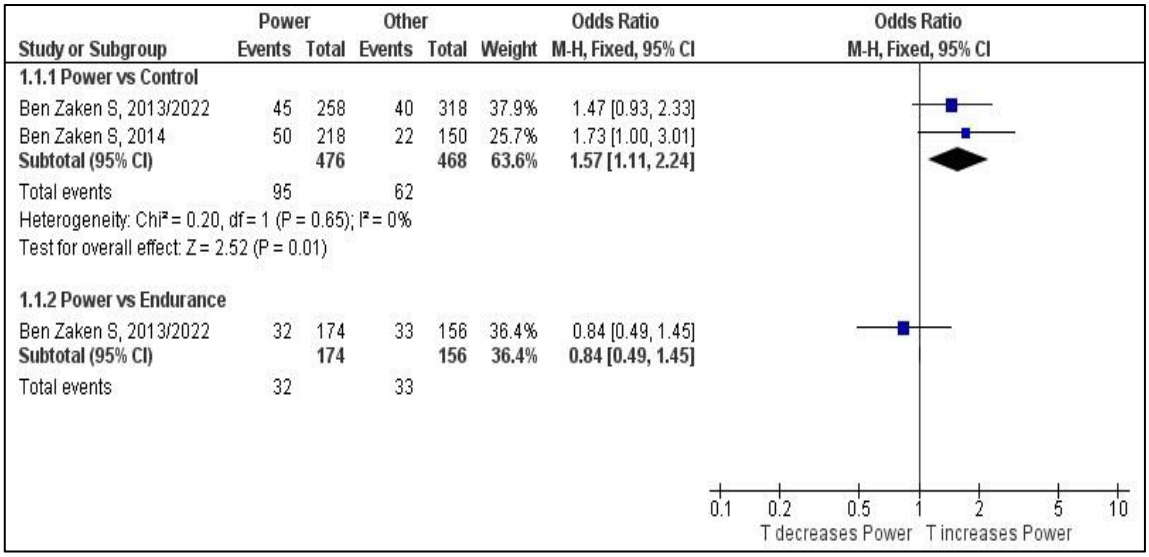

Figure 3 – Forest chart comparing power sport athletes and controls. The results correspond to the merge of Ben-Kazen et al 2013 and Ben-Zaken et al 2014.

Statistically significant differences were highlighted for the frequency of the minor allele between endurance athletes and controls (p= 0.005; OR = 1.91; CI = 1.21-3.01) and between power sport athletes and controls (p = 0.01; OR= 1.57; CI = 1.11-2.24), the minor allele being more frequent in both athletes' groups (Figures 2 and 3). There was no statistically significant difference between power sport and endurance athletes (Figure 3). Hardy-Weinberg equilibrium was verified in controls (after removing controls' duplicates) ($p>0.05$).

Discussion

As the sports phenomena keeps growing more notoriously and moving large economic resources, the demand for finding and developing new athletic talents is at an all-time high, with substantial efforts being poured into all involved areas such as nutrition, sports medicine, conditioning programs, and a recent growing interest in genetics. The variant rs35767 influences the rate of *IGF1* gene transcription and the level of the encoded protein (Palles C 2008; Canzian 2006; Ben-Zaken S. et al. 2015), and has been widely associated with sport-related phenotypes and sport performance (Ben-Zaken S. et al. 2013, 2014, 2022, Karpowicz K. et al. 2018, Ben-Zaken S. et al. 2017, Batavani M.R. et al. 2018, Moreland,E. et al. 2022). As most publications are case-control studies having a limited statistical power due to small samples sizes, we aimed to performed a systematic review and metanalysis to evaluate the association of rs35767 with power and endurance sports performance.

Our results showed a statistically significantly association of the minor allele with endurance and power sport athletes when comparing with non-athletes (p= 0.005; OR = 1.91; CI = 1.21-3.01) and (p = 0.01; OR= 1.57; CI = 1.11-2.24), confirming previous publications (Ben-Zaken S. et al., 2013, Ben-Zaken S. et al., 2022, Ben-Zaken S. et al., 2014). Even though the values of OR obtained are not high, these results are in line with the characteristic low penetrance of genetic variants involved in the etiology of multifactorial, complex, traits, and with their polygenic nature, and are similar to values described even when applying large size populations and genomic methodologies (Ahmetov, I. et al. 2015; Al-Khelaifi, F et al. 2020; Bulgay, C. et al. 2023, Semenova, E.A et al. 2023). Moreover, the regulation of the expression of *IGF1* is complex, with many factors interfering, from the transcription to post-translational level, implying that the effect of rs35767 may be modified (Bailes 2021). We also could not analyze subgroups of elite or national levels athletes, or gender distribution, because of incomplete data available. In fact, in order to identify genetic variants useful to select "talents", it would be more adequate to compare elite athletes with long carrier, national level, athletes that were unable to reach highest standards, than to resort a control group of common individuals.

Association studies aiming to identify genetic variants associated with multifactorial traits, from sport performance to clinical entities, have well known limitations (Varillas-Delgado 2022). Mixed models, including both polygenic risk scores, characterization of the epigenome and exposome, and accounting for variables interaction, may be better suited to predict these phenotypes with sufficient accuracy to make them useful (Odintsova V. et al. 2021).

Notwithstanding, in the future, knowledge on the genetic variants influencing sports performance may be used to refine the already existing methodologies used not only in talent detection, but also in athletes' development and well-being, with impact in the personalization of strength and conditioning programs, nutritional plans, and rehabilitation protocols. Yet, some caution must be taken when allowing genomics to take its place in talented athletes' selection, as legal and ethical questions naturally arise. Can athletes be denied an elite carrier based in their individual genetic profile? Protection of individual genetic information and sanctioning of genetic discrimination is contemplate in many countries' legislation and international institutions (EU Charter of Fundamental Rights, Article 21.1). Translation of genetic profile information to the selection of specific training protocols would otherwise be legitimate and advantageous.

This metanalysis has some limitations: only three studies could be included, the total sample size is still limited, mainly in controls, as controls of Ben-Zaken 2013 and Ben-Zaken 2014 are the same. For statistical analysis and evaluation of Hardy-Weinberg equilibrium in controls this issue was corrected. As the athletes included in each publication practiced different sports, cases were not from the same cohort. Still, we were able to include a total of 354 athletes, thus increasing the statistical power of previous associations of rs35767 with sport performance.

Conclusion

In conclusion, this metanalysis highlighted that the minor allele of rs35767 in *IGF1* gene is more frequent among endurance and power sport athletes comparing with non-athletes, supporting the role of *IGF1* in sport performance and that rs35767 has potential for being a marker of athletes' specific training selection.

References

Ahmetov, I.; Kulemin, N.; Popov, D.; Naumov, V.; Akimov, E.; Bravy, Y.; Egorova, E.; Galeeva, A.; Generozov, E.; Kostryukova,E.; et al. Genome-wide association study identifies three novel genetic markers associated with elite endurance performance. Biol.Sport **2015**, 32, 3–9.

Al-Khelaifi, F.; Yousri, N.A.; Diboun, I.; Semenova, E.A.; Kostryukova, E.S.; Kulemin, N.A.; Borisov, O.V.; Andryushchenko, L.B.; Larin, A.K.; Generozov, E.V.; et al. Genome-Wide Association Study Reveals a Novel Association Between MYBPC3 Gene Polymorphism, Endurance Athlete Status, Aerobic Capacity and Steroid Metabolism. Front. Genet. **2020**, 11, 595.

Bailes, J.; Soloviev, M. Insulin-Like Growth Factor-1 (IGF-1) and Its Monitoring in Medical Diagnostic and in Sports. Biomolecules 2021, 11, 217. https://doi.org/10.3390/biom11020217

Ben Zaken, S., Meckel, Y., Dror, N., Nemet, D., & Eliakim, A. (2014). IGF-I and IGF-I receptor polymorphisms among elite swimmers. Pediatr Exerc Sci, 26(4), 470-476. https://doi.org/10.1123/pes.2014-0158

Ben-Zaken S, Meckel Y, Nemet D, Eliakim A. Can IGF-I polymorphism affect power and endurance athletic performance? Growth Horm IGF Res. 2013 Oct;23(5):175-8. doi: 10.1016/j.ghir.2013.06.005. Epub 2013 Jul 11. PMID: 23850449.

Ben-Zaken S, Meckel Y, Remmel L, Nemet D, Jürimäe J, Eliakim A. The prevalence of IGF-I axis genetic polymorphisms among decathlon athletes. Growth Horm IGF Res. 2022 Jun;64:101468. doi: 10.1016/j.ghir.2022.101468. Epub 2022 May 18. PMID: 35605565.

Ben-Zaken, S., Eliakim, A., Nemet, D., Kaufman, L., & Meckel, Y. (2022). Genetic characteristics of competitive swimmers: a review. *Biology of sport*, *39*(1), 157–170. https://doi.org/10.5114/biolsport.2022.102868

Beunen, G., & Thomis, M. (1999). Genetic determinants of sports participation and daily physical activity. *International Journal of Obesity*, *23*(3), S55–S63. https://doi.org/10.1038/sj.ijo.0800885

Bikle D.D, Tahimic C, Wenhan Chang, Yongmei Wang, Anastassios Philippou, Elisabeth R. Barton, Role of IGF-I signaling in muscle bone interactions, Bone,Volume 80,2015,pag 79-88, https://doi.org/10.1016/j.bone.2015.04.036.

Bouchard, C., Chagnon, M., Thibault, M.-C., Boulay, M. R., Marcotte, M., Cote, C., & Simoneau, J.-A. (1989). Muscle genetic variants and relationship with performance and trainability. Medicine & Science in Sports & Exercise, 21(1), 71-78. https://journals.lww.com/acsm-msse/fulltext/1989/02000/muscle_genetic_variants_and_relationship_with.13.aspx

Boulygina, E.A.; Borisov, O.V.; Valeeva, E.V.; Semenova, E.A.; Kostryukova, E.S.; Kulemin, N.A.; Larin, A.K.; Nabiullina, R.M.;
Mavliev, F.A.; Akhatov, A.M.; et al. Whole genome sequencing of elite athletes. Biol. Sport 2020, 37, 295–304.

Bulgay, C.; Kasakolu, A.; Kazan, H.H.; Mijaica, R.; Zorba, E.; Akman, O.; Bayraktar, I.; Ekmekci, R.; Koncagul, S.; Ulucan, K.; et al.
Exome-Wide Association Study of Competitive Performance in Elite Athletes. Genes 2023, 14, 660.

Cagnin, S., Chemello, F., & Ahmetov, I. I. (2019, January 1). Chapter Eight - Genes and response to aerobic training (D. Barh & I. I. Ahmetov, Eds.). ScienceDirect; Academic VER REFS Press. https://www.sciencedirect.com/science/article/abs/pii/B9780128161937000087

Canzian F, J.D. McKay, R.J. Cleveland, et al Polymorphisms of genes coding for insulin-like growth factor 1 and its major binding proteins, circulating levels of IGF-I and IGFBP-3 and breast cancer risk: results from the EPIC study Br. J. Cancer, 94 (2006), pp. 299-307, 10.1038/sj.bjc.6602936

Cheng, I., Stram, D. O., Penney, K. L., Pike, M., Le Marchand, L., Kolonel, L. N., Hirschhorn, J., Altshuler, D., Henderson, B. E., & Freedman, M. L. (2006). Common genetic variation in IGF1 and prostate cancer risk in the Multiethnic Cohort. *Journal of the National Cancer Institute*, *98*(2), 123–134. https://doi.org/10.1093/jnci/djj013

De Moor MH, Spector TD, Cherkas LF, Falchi M, Hottenga JJ, Boomsma DI, De Geus EJ. Genome-wide linkage scan for athlete status in 700 British female DZ twin pairs. Twin Res Hum Genet. 2007 Dec;10(6):812-20. doi: 10.1375/twin.10.6.812. PMID: 18179392.

Díaz Ramírez, J.; Álvarez-Herms, J.; Castañeda-Babarro, A.; Larruskain, J.; Ramírez de la Piscina, X.; Borisov, O.V. ; Semenova,E.A.; Kostryukova, E.S.; Kulemin, N.A.; Andryushchenko, O.N.; et al. The GALNTL6 Gene rs558129 Polymorphism is Associated with Power Performance. J. Strength Cond. Res. 2020, 34, 3031–3036.

Eriksson, M., Rasmussen, F., & Tynelius, P. (2006). Genetic Factors in Physical Activity and the Equal Environment Assumption--The Swedish Young Male Twins Study. In *Behavior Genetics* (Vol. 36, Issue 2, pp. 238–247). Springer. https://doi.org/10.1007/s10519-005-9018-7

Ernst, S.; Simon, P. A quantitative approach for assessing significant improvements in elite sprint performance: Has IGF-1 entered the arena? *Drug Test. Anal.* 2013, *5*, 384–389

Fink, H. H., & Mikesky, A. E. (2017). Practical Applications in Sports Nutrition (5a ed.).Jones&BartlettLearning,LLC. https://samples.jblearning.com/0763726575/Chapter_12_Fink_Practical_Applications_in_Sports_Nutrition.pdf

García-Arnés JA, García-Casares N. Doping and sports endocrinology: growth hormone, IGF-1, insulin, and erythropoietin, Revista Clínica Española (English Edition), Volume 223, Issue 3, 2023, Pages 181-187, https://doi.org/10.1016/j.rceng.2023.01.005.

Gonzalez E, M.L. Messi, Z. Zheng, O. Delbono, Insulin-like growth factor-1 prevents age-related decrease in specific force and intracellular $Ca2^{+}$ in single intact muscle fibres from transgenic mice, J. Physiol. 552 (Pt 3) (2003) 833–844.

Guilherme, J.P.L.F.; Semenova, E.A.; Zempo, H.; Martins, G.L.; Lancha-Junior, A.H.; Miyamoto-Mikami, E.; Kumagai, H.; Tobina, T.; Shiose, K.; Kakigi, R.; et al. Are Genome-Wide Association Study Identified Single-Nucleotide Polymorphisms Associated with Sprint Athletic Status? A Replication Study With 3 Different Cohorts. Int. J. Sport. Physiol. Perform. 2021, 16, 489–495

Gusscott, S.; Jenkins, C.E.; Lam, S.H.; Giambra, V.; Pollak, M.; Weng, A.P. IGF1R derived PI3K/AKT signaling maintains growth in a subset of human T-cell acute lymphoblastic leukemias. PLoS ONE **2016**, 11, e0161158.

Hernandez, W., Grenade, C., Santos, E. R., Bonilla, C., Ahaghotu, C., & Kittles, R. A. (2007). IGF-1 and IGFBP-3 gene variants influence on serum levels and prostate cancer risk in African-Americans. Carcinogenesis, 28(10), 2154–2159. https://doi.org/10.1093/carcin/bgm190

Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. BMJ (Clinical research ed.), 327(7414), 557–560. https://doi.org/10.1136/bmj.327.7414.557

Kashimoto, R. K., Toffoli, L. V., Manfredo, M. H. F., Volpini, V. L., Martins-Pinge, M. C., Pelosi, G. G., et al. (2016). Physical Exercise Affects the Epigenetic Programming of Rat Brain and Modulates the Adaptive Response Evoked by Repeated Restraint Stress. Behav. Brain Res. 296, 286–289. doi:10.1016/ j.bbr.2015.08.038

Kostek M.C. , J.M. Devaney, H. Gordish-Dressman, et al., A polymorphism near IGF1 is associated with body composition and muscle function in women from the Health, Aging, and Body Composition Study, Eur. J. Appl. Physiol. 110 (2) (2010) 315–324.

Koziris LP, Hickson RC, Chatterton RT Jr et al (1999) Serum levels of total and free IGF-I and IGFBP-3 are increased and maintained in long-term training. J Appl Physiol 86:1436–1442

Laron Z. Insulin-like growth factor 1 (IGF-1): a growth hormone. Mol Pathol. 2001 Oct;54(5):311-6. doi: 10.1136/mp.54.5.311. PMID: 11577173; PMCID: PMC1187088.

Lippi, G., Longo, U. G., & Maffulli, N. (2010). Genetics and sports. *British Medical Bulletin*, *93*(1), 27–47. https://doi.org/10.1093/bmb/ldp007

Ma F, Yang Y, Li X, Zhou F, Gao C, Li M, et al. (2013) The Association of Sport Performance with *ACE* and *ACTN3* Genetic Polymorphisms: A Systematic Review and Meta-Analysis. PLoS ONE 8(1): e54685. https://doi.org/10.1371/journal.pone.0054685

Maia, J. A. R., Thomis, M., & Beunen, G. (2002). Genetic factors in physical activity levels: a twin study. *American Journal of Preventive Medicine*, *23*(2 Suppl), 87–91. https://doi.org/10.1016/s0749-3797(02)00478-6

Maïmoun, L., Galy, O., Manetta, J., Coste, O., Peruchon, E., Micallef, J. P., Mariano-Goulart, D., Couret, I., Sultan, C., & Rossi, M. (2004). Competitive season of triathlon does not alter bone metabolism and bone mineral status in male triathletes. *Int J Sports Med*, *25*(3), 230-234. https://doi.org/10.1055/s-2003-45257

Malczewska-Lenczowska, J.; Orysiak, J.; Majorczyk, E.; Sitkowski, D.; Starczewski, M.; Zmijewski, P. HIF-1α and NFIA-AS2
Polymorphisms as Potential Determinants of Total Hemoglobin Mass in Endurance Athletes. J. Strength Cond. Res. 2022, 36,1596–1604.

Marx, J. O., Ratamess, N. A., Nindl, B. C., Gotshalk, L. A., Volek, J. S., Dohi, K., Bush, J. A., Gómez, A. L., Mazzetti, S. A., Fleck, S. J., Häkkinen, K., Newton, R. U., & Kraemer, W. J. (2001). Low-volume circuit versus high-volume periodized resistance training in women. Med Sci Sports Exerc, 33(4), 635-643. https://doi.org/10.1097/00005768-200104000-00019

Miyamoto-Mikami, E., Zempo, H., Fuku, N., Kikuchi, N., Miyachi, M., & Murakami, H. (2018). Heritability estimates of endurance-related phenotypes: A systematic review and meta-analysis. *Scandinavian journal of medicine & science in sports*, *28*(3), 834–845. https://doi.org/10.1111/sms.12958

McAuley, A. B. T., Hughes, D. C., Tsaprouni, L. G., Varley, I., Suraci, B., Roos, T. R., Herbert, A. J., & Kelly, A. L. (2021). The association of the *ACTN3* R577X and *ACE* I/D polymorphisms with athlete status in football: a systematic review and meta-analysis. *Journal of sports sciences*, *39*(2), 200–211. https://doi.org/10.1080/02640414.2020.1812195

Odintsova VV, Rebattu V, Hagenbeek FA, Pool R, Beck JJ, Ehli EA, van Beijsterveldt CEM, Ligthart L, Willemsen G, de Geus EJC, Hottenga JJ, Boomsma DI, van Dongen J. Predicting Complex Traits and Exposures From Polygenic Scores and Blood and Buccal DNA Methylation Profiles. Front Psychiatry. 2021 Jul 29;12:688464. doi: 10.3389/fpsyt.2021.688464. PMID: 34393852; PMCID: PMC8357987.

Ostrovsky O, Eletto D, C. Makarewich, E.R. Barton, Y. Argon Glucose regulated protein 94 is required for muscle differentiation through its control of the autocrine production of insulin-like growth factors Biochim Biophys Acta, 1803 (2010), pp. 333-341

Page, M.J. et al. (2021) Prisma 2020 explanation and elaboration: Updated guidance and exemplars for Reporting systematic reviews, The BMJ. Available at: https://www.bmj.com/content/372/bmj.n160?gclid=Cj0KCQjw1rqkBhCTARIsAAHz7K11QGpod19pMuiNu0IkgFYUvUdAsaUkhLNsiP9voR92jiC-7zR4vGkaAuJ_EALw_wcB (Accessed: 21 August 2023).

Palles C, N. Johnson, B. Coupland, et al. Identification of genetic variants that influence circulating IGF1 levels: a targeted search strategy Hum. Mol. Genet., 17 (2008), pp. 1457-1464, 10.1093/hmg/ddn034

Pei, YF., Liu, YZ., Yang, XL. et al. The genetic architecture of appendicular lean mass characterized by association analysis in the UK Biobank study. Commun Biol 3, 608 (2020). https://doi.org/10.1038/s42003-020-01334-0

Pickering, C.; Suraci, B.; Semenova, E.A.; Boulygina, E.A.; Kostryukova, E.S.; Kulemin, N.A.; Borisov, O.V.; Khabibova, S.A.; Larin,A.K.; Pavlenko, A.V.; et al. A genome-wide association study of sprint performance in elite youth football players. J. Strength Cond. Res. 2019, 33, 2344–2351.

Pittaluga, M., Casini, B., & Parisi, P. (2004). Physical activity and genetic influences in risk factors and aging: a study on twins. *International Journal of Sports Medicine*, *25*(5), 345–350. https://doi.org/10.1055/s-2004-815847

Pitsiladis YP, Tanaka M, Eynon N, Bouchard C, North KN, Williams AG, et al. Athlome project consortium: A concerted effort to discover genomic and other "omic" markers of athletic performance. Physiol Genomics. 2016;48:183–90.

Rankinen, T.; Fuku, N.; Wolfarth, B.; Wang, G.; Sarzynski, M.A.; Alexeev, D.G.; Ahmetov, I.I.; Boulay, M.R.; Cieszczyk, P.; Eynon,
N.; et al. No evidence of a common DNA variant profile specific to world class endurance athletes. PLoS ONE 2016, 11, e0147330.

Regateiro, F. (2007). *Manual de Genética médica*. Universidade de Coimbra

Roelen CA, de Vries WR, Koppeschaar HP, Vervoorn C, Thijssen JH, Blankenstein MA (1997) Plasma insulin-like growth factor-I and high affinity growth hormone-binding protein levels increase after two weeks of strenuous physical training. Int J Sports Med 18(4):238–241

Rubin MR, Kraemer WJ, Maresh CM et al (2005) High-affinity growth hormone binding protein and acute heavy resistance exercise. Med Sci Sports Exerc 37:395–403

Seaborne, R. A., and Sharples, A. P. (2020). The Interplay between Exercise Metabolism, Epigenetics, and Skeletal Muscle Remodeling. Exerc. Sport Sci. Rev. 48 (4), 188–200. doi:10.1249/JES.0000000000000227

Sellami M, Elrayess MA, Puce L and Bragazzi NL (2022) Molecular Big Data in Sports Sciences: State-of-Art and Future Prospects of OMICS-Based Sports Sciences. Front. Mol. Biosci. 8:815410. doi: 10.3389/fmolb.2021.815410

Semenova, E.A.; Zempo, H.; Miyamoto-Mikami, E.; Kumagai, H.; Larin, A.K.; Sultanov, R.I.; Babalyan, K.A.; Zhelankin, A.V.;Tobina, T.; Shiose, K.; et al. Genome-Wide Association Study Identifies CDKN1A as a Novel Locus Associated with Muscle Fiber Composition. Cells 2022, 11, 3910.

Semenova, E. A., Hall, E. C. R., & Ahmetov, I. I. (2023). Genes and Athletic Performance:The 2023 Update. Genes, 14(6), 1235. https://doi.org/10.3390/genes14061235

Semenova, E. A., Pranckevičienė, E., Bondareva, E. A., Gabdrakhmanova, L. J., & Ahmetov, I. I. (2023). Identification and Characterization of Genomic Predictors of Sarcopenia and Sarcopenic Obesity Using UK Biobank Data. Nutrients, 15(3), 758. https://doi.org/10.3390/nu15030758

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. https://doi.org/10.1080/02640410903110974

Varillas-Delgado, D., Del Coso, J., Gutiérrez-Hellín, J., Aguilar-Navarro, M., Muñoz, A., Maestro, A., & Morencos, E. (2022). Genetics and sports performance: the present and future in the identification of talent for sports based on DNA testing. European journal of applied physiology, 122(8), 1811–1830. https://doi.org/10.1007/s00421-022-04945-z

Vaughan D, Brogioli M, Maier T, White A, Waldron S, Rittweger J, et al. (2016) The Angiotensin Converting Enzyme Insertion/Deletion Polymorphism Modifies Exercise-Induced Muscle Metabolism. PLoS ONE 11(3): e0149046. https://doi.org/10.1371/journal.pone.0149046

Velloso CP. (2008) **:** Regulation of Muscle Mass by Growth Hormone and IGF-I. Br J Pharmacol 154(3):557–68.

Veuille M. (March 23, 2019), Genetic polymorphism and variation, Encyclopedia of the Environment, Accessed August 21, 2023 [online ISSN 2555-0950] url : https://www.encyclopedie-environnement.org/en/life/genetic-polymorphism-and-variation/.

Yakar S, Werner H, Rosen CJ. (2018) : Insulin-Like Growth Factors: Actions on the Skeleton. J Mol Endocrinol 61(1):T115–t37.

Yengo, L., Vedantam, S., Marouli, E. et al. A saturated map of common genetic variants associated with human height. Nature 610, 704–712 (2022). https://doi.org/10.1038/s41586-022-05275-y

Wang Y, Cheng Z, H.Z. Elalieh, E. Nakamura, M.T. Nguyen, S. Mackem, *et al.* IGF-1R signaling in chondrocytes modulates growth plate development by interacting with the PTHrP/Ihh pathway J Bone Miner Res, 26 (2011), pp. 1437-1446

Webborn N, Williams A, McNamee M, Bouchard C, Pitsiladis Y, Ahmetov I, et al. Direct-to-consumer genetic testing for predicting sports performance and talent identification: Consensus statement. Br J Sports Med. 2015;49:1486–91.

Zehra, Z., Khan, N., Nadeem, M., Siddiqui, S. N., von Bartheld, C. S., Azam, M., & Qamar, R. (2022). Association of *IGF1* polymorphisms with exotropia in a Pakistani cohort. *Molecular vision*, *28*, 369–377.

Appendix

Appendix A

Newcasttle-Ottawa Scale (NOS) for Assessing risk of Bias
NCBI Bookshelf. A service of the National Library of Medicine, National Institutes of Health.

Gierisch JM, Beadles C, Shapiro A, et al. Health Disparities in Quality Indicators of Healthcare Among Adults with Mental Illness [Internet]. Washington (DC): Department of Veterans Affairs (US); 2014 Oct.

NEWCASTLE-OTTAWA SCALE CODING MANUAL FOR COHORT STUDIES

The Newcastle-Ottawa Scale quality instrument is scored by awarding a point for each answer that is marked with an asterisk below. Possible total points are 4 points for Selection, 2 points for Comparability, and 3 points for Outcomes.

SELECTION

1. Representativeness of the Exposed Cohort

a. Truly representative of the average patient with mental illness (eg, severity of illness, comorbidities) in the community*

b. Somewhat representative of the average (eg, severity of illness, comorbidities)in the community*

c. Selected group of users eg HIV+, pregnant, elderly, significant physical disabilities

d. No description of the derivation of the cohort

2. Selection of the Non-Exposed Cohort

a. Drawn from the same community as the exposed cohort*

b. Drawn from a different source

c. No description of the derivation of the non-exposed cohort

3. Ascertainment of Exposure

a. Secure record (eg, medical records)*

b. Structured interview *

c. Written self-report

d. No description

4. Demonstration that Outcome of Interest Was Not Present at Start of Study

a. Yes*

b. No

COMPARABILITY

1. Comparability of Cohorts on the Basis of the Design or Analysis

a. Study controls for SES (or some reasonable proxy of SES), age, race, gender*

b. Study controls for any additional factor* (this criteria could be modified to indicate specific control for a second important factor)

c. Inadequate degree of control

OUTCOME

1. Assessment of Outcome

a. Independent or blind assessment stated in the paper, or confirmation of the outcome by reference to secure records (x-rays, medical records, etc)*

b. Record linkage (eg, identified through ICD codes on database records)*

c. Self-report (ie, no reference to original medical records or x-rays to confirm the outcome)

d. No description

2. Was Follow-up Long Enough for Outcomes to Occur?

a. Yes (select an adequate follow up period for outcome of interest)*

b. No

3. Adequacy of Follow-up of Cohorts

a. Complete follow-up—all subjects accounted for*

b. Subjects lost to follow-up unlikely to introduce bias—small number lost (LESS than 20% follow-up, or description provided of those lost)*

c. Follow-up rate MORE than 20% and no description of those lost

d. No statement