Title

Doping prevalence in competitive sport: Evidence synthesis with "best practice" recommendations and reporting guidelines from the WADA Working Group on Doping Prevalence

Authors

John Gleaves[1,*], Andrea Petróczi[2], Dirk Folkerts[2,3], Olivier de Hon[4], Emmanuel Macedo[5], Martial Saugy[6]

[1] California State University Fullerton

[2] Kingston University London

[3] University of Münster

[4] Doping Authority Netherlands

[5] Penn State University

[6] University of Lausanne

* Corresponding author

Contact address: John Gleaves, Department of Kinesiology, California State University Fullerton, 800 N State College Blvd, Fullerton, CA 92834. jgleaves@fullerton.edu

Abstract

Background

The prevalence of doping in competitive sport, and the methods for assessing prevalence, remain poorly understood. This reduces the ability of researchers, governments, and sporting organizations to determine the extent of doping behavior and the impacts of anti-doping strategies.

Objectives

The primary aim of this subject- wide systematic review was to collate and synthesize evidence on doping prevalence from published scientific papers. Secondary aims involved reviewing the reporting accuracy and data quality as evidence for doping behavior to (1) develop quality and bias assessment criteria to facilitate future systematic reviews; and (2) establish recommendations for reporting future research on doping behavior in competitive sports to facilitate better meta-analyses of doping behavior.

Methods

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines were used to identify relevant studies. Articles were included if they contained information on doping prevalence of any kind in competitive sport, regardless of the methodology and without time limit. Through an iterative process, we simultaneously developed a set of assessment criteria; and used these to assess the studies for data quality on doping prevalence, potential bias and reporting.

Results

One-hundred and five studies, published between 1975 and 2019,were included. Doping prevalence rates in competitive sport ranged from 0% to 73% for doping behavior with most falling under 5%. To determine prevalence, 89 studies used self-reported survey data (SRP) and 17 used sample analysis data (SAP) to produce evidence for doping prevalence (one study used both SRP and SAP). In total, studies reporting athletes totaled 102,515 participants, (72.8% men and 27.2% women). Studies surveyed athletes in 35 countries with 26 involving athletes in the United States while 12 studies examined an international population. Studies also surveyed athletes from most international sport federations and major professional sports and examined international, national, and sub-elite level athletes, including youth, masters, amateur, club, and university level athletes. However, inconsistencies in data reporting prevented meta-analysis for sport, gender, region, or competition level. Qualitative syntheses were possible and provided for study type, gender, and geographical region.

The quality assessment of prevalence evidence in the studies identified 20 as “High,” 60 as “Moderate,” and 25 as “Low.” Of the 89 studies using SRP, 17 rated as “High,” 52 rated as “Moderate,” and 20 rated as “Low.” Of the 17 studies using SAP, 3 rated as “High,” 9 rated as “Moderate, and 5 rated as “Low.” Examining ratings by year suggests that both the quality and quantity of the evidence for doping prevalence in published studies are increasing.

Conclusions

Current knowledge about doping prevalence in competitive sport relies upon weak and disparate evidence. To address this, we offer a comprehensive set of assessment criteria for studies examining doping behavior data as evidence for doping prevalence. To facilitate future evidence

syntheses and meta-analyses, we also put forward “best practice” recommendations and reporting guidelines that will improve evidence quality.

Key points

- Reported or estimated doping in competitive sport prevalence rates in all studies ranged between 0% and 73%, with most falling under 5%.
- Studies surveyed totaled 102,515 participants (72.8% men and 27.2% women) sampled between 1976 to 2019 from over 35 countries with 12 studies including an international population.
- Self-reports on doping behavior in anonymous surveys comprise 81.7% of the literature on doping prevalence.
- Evidence for the prevalence of doping in competitive sport remains fragmented due to inconsistent study design and reporting.
- “Best practice” recommendations and reporting guidelines may improve the quality of evidence for doping prevalence.

1. INTRODUCTION

1.1. Background

Governments and sporting organizations are under pressure to prevent doping, whereby athletes intentionally use prohibited substances to enhance performances in competitive sports.[1] By 2014, such pressure had led to an overall spending on anti-doping to approach US$500 million [2], with US$35 million spent directly by the World Anti-Doping Agency [3]. Without reliable estimates for doping prevalence, the effects of such efforts to reduce doping use in competitive sport remains unknown.

Critics urging doping prevalence estimates [4, 5], perhaps inadvertently, create the impression that doping prevalence is easy to assess and that the absence of such figures stems from the reluctance of anti-doping administrators, such as the World Anti-Doping Agency (WADA), to generate potentially unflattering numbers. Yet rather than political calculations, significant methodological challenges may explain neither why anti-doping agencies nor researchers have previously offered scientifically sound estimates of doping among competitive athletes.

A primary obstacle for determining doping prevalence stems from doping not only being against the rules of competitive sport but also severely punished, socially stigmatized, and often illegal.[2] As such, research is showing that athletes who dope are increasingly unwilling to disclose their activities to anyone leaving teammates, family, and support personnel unaware of such activities

[1] There is not a universal definition of doping. However, this study builds upon [1] definition where doping "refers to the set of prohibited substances and/or methods as identified by the ruling body of the particular sport," which, "means that the term 'doping' in […] does not reflect other doping violations mentioned in the World Anti-Doping Code, such as whereabouts failures or trafficking." We have also differentiated between therapeutic and unintentional use of prohibited substances to more clearly describe the phenomenon.

[2] The connection between controlled substances in sport (doping) and in general is a complicated one. First of all, not all substances prohibited in sport is a controlled substance for the general population, and it varies from one country to another. One example for this is anabolic steroids (AS). AS is prohibited in sport both in- and out-of-competition for all athletes around the globe under WADA regulations. However, whilst using AS is also illegal in some countries (e.g., Australia, US, Norway, Saudi Arabia), in other countries (e.g., UK, Canada, South Africa, Turkey) personal use is not illegal but production and supply without license are, regardless of who uses it. In countries where doping is a criminal offence (e.g., Austria, Germany, France, Italy, Israel), AS use is only illegal and can carry a prison sentence for athletes if they are subject to doping control, but not for the general population. AS is not a controlled substance in some countries (e.g., Japan, Bulgaria, Russia, Mexico).

[6]. Even with the promise of anonymity, athletes seem unlikely to admit to doping when surveyed by an unknown researcher. Additionally, methods for doping and evasion of doping testing are constantly evolving. Athletes who dope often go to great lengths to avoid detection, which limits the reliability of testing blood and urine samples as a measure of prevalence. Compounding these issues, the nuances of anti-doping rules easily allow poorly worded surveys to generate misleading estimates. For example, many substances prohibited by anti-doping rules also have therapeutic benefits such that an athlete may be taking a banned substance under a therapeutic use exemption (TUE), which permits athletes to use prohibited substances necessary to treat medically validated conditions (e.g. dextroamphetamine/amphetaminefor attention-deficit/hyperactivity disorder). Moreover, some substances are banned only for certain sports or for use during competition (e.g. beta-blockers in certain sports). A poorly worded survey or poor explanation of anti-doping rules may lead respondents to indicate doping activities despite never having done so.

Finally, prevalence estimates always reflect a defined population. Defining a population presents a challenge for doping prevalence in competitive sport where the populations can be fluid and diffuse. The population of professional football players may change significantly from season to season while the population of “elite” level athletes may be unclear because there is no rigid definition for when an athlete has actually become “elite.” At the same time, the diffuse nature of sport means that a survey of German triathletes may not say much about their Japanese counterparts or even triathletes as a whole. Such difficulties are demonstrated in a review on doping prevalence in New Zealand [7], which illustrates the challenges to generating prevalence numbers for specific sports when considering who counts as a member of the population.

1.2. WADA Working Group on Doping Prevalence

In 2017, WADA reconvened an expert working group on doping prevalence. The working group's mission was to establish a better understanding of the prevalence of doping in competitive sport. The members of the Working Group on Doping Prevalence (AP, JG, MS and OdH) were internationally recognized experts with scholarly backgrounds in doping research. The working group determined that a systematic review and evidence synthesis of doping prevalence would be a necessary first step for its purposes and could potentially benefit the scholarly community researching doping prevalence based upon the following rationale.

1.3. Rationale

Having reliable information on the extent of doping use in competitive (and thus regulated) sports is paramount for devising appropriate doping control and prevention programs. In the literature, a limited number of reviews on prevalence (mainly focusing on methodological issues) present some insights into this hidden practice. Unfortunately, these review studies fail to offer a definitive picture of doping prevalence in competitive. For example, a review by Dimeo and Taylor [8] does not follow a systematic method for identifying relevant studies and mixes prevalence reports studies of doping attitude and implicit associations with perception of doping prevalence and doping intentions. Although social science literature often uses such measures as a proxy for doping behavior, they cannot be interpreted as prevalence figures because they reveal respondents' beliefs rather than actual practices within a population [9]. The review by de Hon et al. [1] two years later uses a systematic search but includes survey studies with fitness center visitors. Both reviews also include studies with amateur athletes, students, exercisers, and gym goers often without a clearly stated distinction.

One significant challenge for all of these reviews remains the fragmented and patchy scholarship on doping prevalence in competitive sport. Studies vary greatly in design and generalizability. This prevents authors from preparing a meaningful systematic review, whereby multiple studies could be pooled and analyzed together to better determine what is known about doping prevalence. The variations in study design also make it difficult to determine which doping prevalence data is of better quality and which is of lesser quality. Researchers unfamiliar with doping prevalence may not know which studies report higher quality evidence or which study methods are more reliable indicators for doping prevalence.

More recent systematic examinations [10-12] focusing on predictors of doping intentions, susceptibility, and behavior of elite athletes identified fourteen studies. In these studies, the presence of doping behavior in the sample was established with self-reported use of doping. Unfortunately, the results from these reviews are confounded by the authors including studies that used self-efficacy and perceived personal control measures and studies of athletes using drugs other than prohibited performance enhancing substances such as illicit recreational drugs or nutritional supplements.

In such cases, a growing consensus [13, 14] supports using a subject-wide evidence synthesis as a valuable alternative capable of providing "a rigorous way to synthesize information when data are unevenly or thinly distributed, or highly variable in focus" [13]. In this case, a subject-wide evidence synthesis provides insights into open-framed questions such as how many studies have reported doping prevalence for their respective samples and what methods such studies used. The answers to such questions can improve decision-makers' and researchers' understanding of doping prevalence by providing access to all available evidence on the issue in question. A comprehensive, accurate and unbiased synthesis of all available evidence in a concise format is

therefore one of the most valuable contributions the research community can offer to inform policymakers and stakeholders.

1.4. Objectives

The aim of this evidence synthesis was twofold. First, the research team set out to provide a systemic mapping [13] of the available evidence in the literature on doping prevalence in competitive sport. Second, the research team intended to assess the evidence quality for doping prevalence with the view of informing future empirical studies investigating doping prevalence or reporting data on doping behavior that can be extracted and pooled in a meta-analysis to establish prevalence. For the latter, investigators sought a set of quality assessment criteria to facilitate better research and reporting.

2. METHODS

2.1. Eligibility criteria

Empirical studies that provided the doping prevalence as a percentage of participants or samples, or studies that provided evidence that showed a doping prevalence could have been calculated for its participants or samples were considered for inclusion regardless of the main purpose of the study.

Publications focusing on population other than competitive athletes (e.g. general populations, exercisers, bodybuilders, university students and pupils who never competed beyond their own school) were excluded. Studies using purposive sampling for doping use (e.g. 50% users, 50% non-users) were also excluded.

Studies focusing only on competitive athletes using substances other than prohibited performance-enhancing drugs were excluded. These substances include the use of illicit (recreational) drugs, nutritional/dietary supplements, prescription medication with Therapeutic Use Exemptions (TUE) or non-prohibited, and non-prohibited over-the counter medication.

Studies reporting "prevalence" based on attitude, susceptibility, intention or other proxy measures were excluded. Unless the prevalence rate was calculated from a concrete number of known users within a personal network and the personal network size (e.g., 3 users known to the respondent from his/her personal network of 24 athletes, giving 12.5% for prevalence in his/her social network), as used in the Network Scale-Up method [15], data from projections (athletes guessing the percentage of other athletes using doping) was also excluded. We excluded such data because when the respondents have no true knowledge of what others do in the subpopulation (e.g., teammates, athletes in the same sport, same country or different country) the responses tend to be influenced by the so called "False Consensus" effect or "Uniqueness Bias" [16], and thus do not offer objective information on doping prevalence.

2.2. Information Sources

The study followed the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) [17], depicted in Figure 1, as a guideline for study identification, selection, inclusion and reporting.

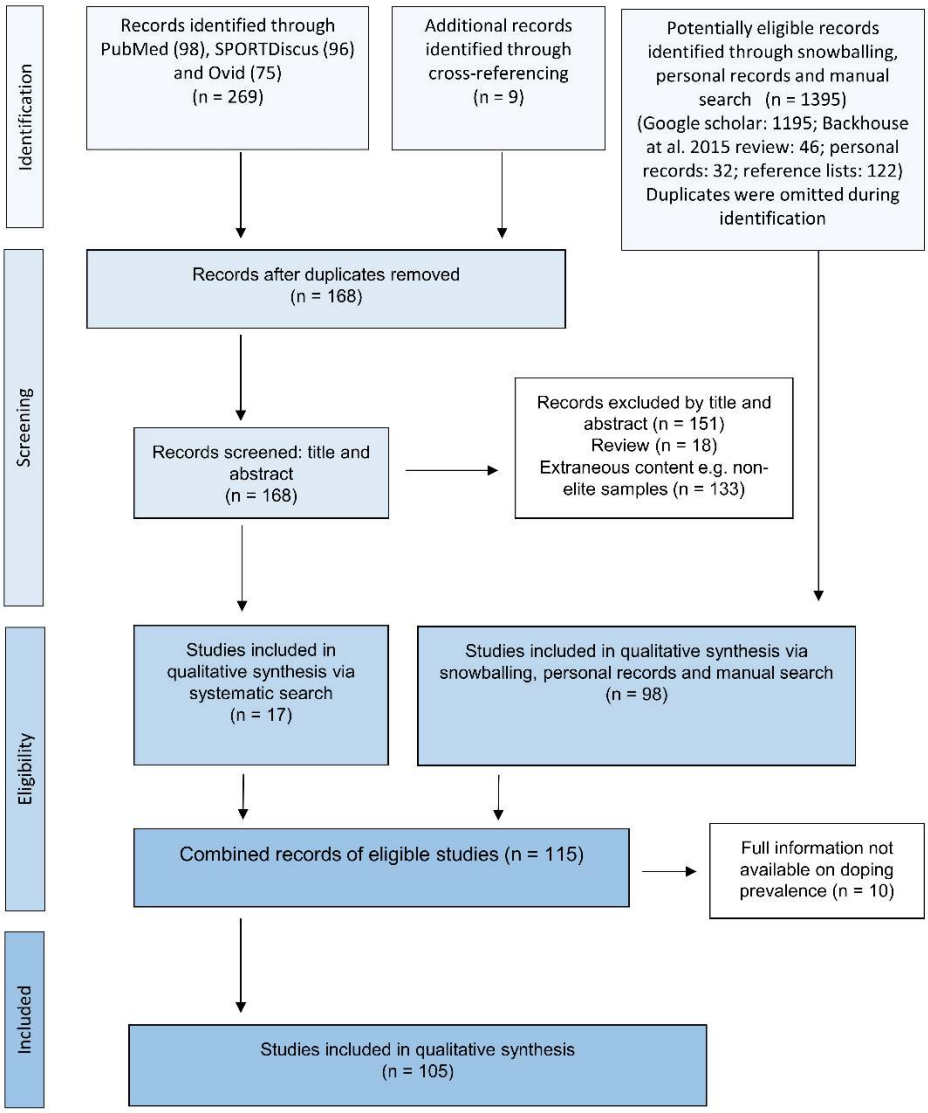

Figure 1: *PRISMA flowchart for doping prevalence studies (without time limit)*

2.3. Search Strategy

The study identified three databases most likely to contain relevant studies: PubMed, SPORTDiscus and Ovid. Additionally, a manual search was performed in Google Scholar. To identify studies which contains information on doping behavior for the sample, the following search terms and combinations were used: “athlete OR player” AND doping AND “TPB OR model” AND sport AND “survey OR questionnaire” AND “report OR admit OR indicate” AND

method AND results AND WADA. This search resulted in 1320 hits excluding citations. Adding "illegal OR illicit OR prohibited OR banned" with an AND operator only reduced the number of hits marginally, to 1170 hits.

Identified results of the literature search were processed and scanned using web application program "Rayyan" [18]. Findings of the mentioned databases were extracted as .xml or .ris files enabling processability for Rayyan. The resulting comprehensive dataset was scanned by title and abstract for eligibility. Data extraction and quality assessment were performed by OdH, DF and EM.

Studies cited in de Hon et al. [1] and Dimeo & Taylor [8] were also scanned manually for eligible studies. Additional studies were identified by means of the snowball search technique (i.e., going through the references of studies already found). The latter included the meta-analyses by Ntoumanis *et al.* [12], Blank *et al.* [10], Sagoe et al. [19], and by Backhouse *et al.* [11]. These reviews also contain studies that establish doping behavior for competitive athletes and were manually scanned for inclusion.

We also used expert knowledge of the WADA Working Group on Doping Prevalence to identify potential studies. This combined technique is recommended for reviews of complex evidence such as doping prevalence. Greenhalgh and Peacock [20] show that in reviewing problems with complex evidence, reviews that rely solely on protocol-driven search strategies identify about 30% of the relevant studies, with 50% coming from snowballing and a further 20% through personal knowledge and contacts. Our study conforms to this pattern.

2.4. Data items

Given the existing state of doping prevalence research, we did not seek to provide a single estimate of doping prevalence, even for a specific population or time frame. Instead, we prioritized mapping and synthesizing the diversity of evidence that indicates doping behavior among athletes in competitive sport. To better portray the evidence, we included any study capable of providing some evidence of doping behavior for a defined population. The evidence synthesis then focused on reporting the methods and range of prevalence for the study types.

2.5. Quality Assessment of Doping Evidence

In the absence of a suitable tool for assessing data quality and bias for this evidence synthesis, authors developed two specific tools for assessing the reported evidence for doping prevalence based on the methods used in the study. Because the majority of the evidence was found in studies not specifically set out to establish prevalence, using the bespoke quality assessment tool for prevalence studies [21] was not appropriate. The assessment tool for surveys with self-reported doping behavior combines the Quality Assessment Tool for Systematic Reviews of Observational Studies (QATSO) items and scoring [22] with items from the assessment tool used in Ntoumanis *et al.*'s [12] systematic review and meta-analysis on doping behavior. The final 17 items of the Quality Assessment of Doping evidence – Self-Reported Prevalence (QUAD-SRP) and rater instructions are included in Electronic Supplementary Material Appendix S1. The research team also developed an assessment tool for studies analyzing samples for prohibited substances. The final 8 items of the Quality Assessment of Doping evidence – Sample Analysis of Prevalence (QUAD-SAP) and rater instructions are included in Electronic Supplementary Material Appendix S2.

The included articles were read in full and quality was independently rated by three of the authors (DF, EM, and OdH.). The QUAD-SRP and the QUAD-SAP were both tested on a random set of articles with three raters to establish interrater reliability for the tools. Each rater assessed a random set of 70 studies, ensuring that two raters assessed every study. In the case of discrepancies between raters, consensus was achieved by discussion between the authors who supervised the raters (AP and JG).

The QUAD-SRP and QUAD-SAP scores are calculated by dividing the sum scores for each question by the total number of applicable items. Questions which did not apply to a specific study, such as the requirements for 'Randomized Response Techniques', were rated as not applicable (NA) and omitted from the total number of applicable items. The authors considered both a weighted and unweighted scoring but little difference emerged from weighting thus the final rating used the unweighted score. After testing for interrater reliability, but prior to completing the scoring of the studies, authors applied customary use of quality grading based on nominal quartile ranges of the maximum possible score (100%) to establish a qualitative grades as: a score > 75% of the maximum possible is "High" (green); a score between ≤ 75% and ≥50% is "Moderate" (yellow); and a score < 50% is "Low" (red) quality evidence for doping prevalence to support the evidence synthesis and assist researchers wishing to identify higher quality evidence for doping prevalence. We collapsed the two bottom grades ('very low' < 25% and 'low' < 50%) into "Low" because distinguishing between "Low" and "Very Low" is practically irrelevant. Because the quality assessment scores were calculated based on a model ideal scenario, the qualitative categories through quality grades reflects absolute (criterion-driven), not relative (within sample) quality. In line with the aims of systematic mapping for evidence synthesis, no studies were excluded based on data quality, bias and/or reporting flaws.

3. RESULTS

3.1. Results of the Search

The review identified 115 studies that met the inclusion criteria. Ten studies were excluded from the final set because the study design did not allow for prevalence to be calculated or the study presented data in a language other than English. Seven studies meeting the inclusion criteria were included even though no prevalence data was reported because the study design allowed for the calculating of prevalence but the data was omitted from the published materials. The full table of results is in Electronic Supplementary Material Appendix S3.

After reviewing articles for inclusion, the authors divided the studies into two groups based upon its method used to determine prevalence: those using self-reported (SRP) data and those using sample analysis data (SAP). Such separation was called for because of the different quality assessment criteria for surveys and sample analyses. Studies placed in SRP employed four different methods for determining prevalence, which were Direct Survey (DS) [23-101], Random Response Technique (RT) [102-110], Qualitative Interviews (QI) [111], and Network Scale Up (NS) [55]. Studies placed in SAP employed four different methods for determining prevalence, which were Testing Figures (TF) [112-121], Blood Profile (BP) [122-124], Anti-Doping Rule Violations (AD) [125-127], and Hair Sample (HS) [79]. One study [79] was found to use a method assigned to SRP and to SAP, thus it was included in both with its two assessment scores included independently in Electronic Supplementary Material Appendix S4. Additionally, two studies were found to use two distinct SRP methods to establish prevalence [55, 88] and both of their respective SRP methods were included in the methods count. This provided the review with a total of 108 methods from 105 studies, with James [55], Petróczi [79], and Striegel et al. [88] having each used two methods. The counts for each method are included in Table 1.

3.2. Sources of Evidence for Doping Prevalence

Of the 105 studies, all studies used either surveys to establish self-reported prevalence (SRP) and/or analyzed samples for prohibited substances to establish sample analysis Prevalence (SAP).

Study Group	Study Method	No of methods (N=108)	Doping Prevalence Range	References
SRP	DS	79	0-66.7%	23-101
SRP	RT	10	3.2-57.1%	102-110
SAP	TF	10	0-6.6%	112-121
SAP	BP	3	0-48%	122-124
SAP	AD	3	0.4-2.6%	125-127
SRP	QI	1	0%	111
SAP	HS	1	13.4%	79
SRP	NS	1	19.9-58.4%	55

Table 1. *Numbers of each method employed to establish doping behavior (Three studies [55, 79, 88] were counted twice because multiple methods were used and their prevalence figures were included in the prevalence range for both study methods used in their respective studies). Self-reported prevalence (SRP); Sample analysis for prevalence (SAP); Direct survey (DS); Random response technique (RT); Testing figure (TF); Athletes biological passport (BP); Anti-doping rule violation (AD); Qualitative interview (QI); Hair sample analysis (HS); Network scale-up (NS).*

The vast majority of these studies did not explicitly identify doping prevalence as an intended aim for the study, yet provided information that shed light onto athletes' doping behavior in competitive sport. As the purpose of this evidence synthesis was to be comprehensive, these studies were considered because they provide some evidence on doping prevalence for the sampled population. Moreover, the findings from both the non-prevalence studies and the lower quality prevalence studies have been cited as evidence for doping prevalence, which further supports their inclusion the evidence synthesis.

Table 2 offers an overview of the amount of evidence for doping prevalence ranges. Owing to the significant variations in substances included, timeframe and methodologies, exclusive categorization was not possible. Readers are advised to consults the summary of evidence in Supplementary Material Appendix S3 and the quality assessment in Supplementary Material Appendix S4 as well as the original articles, particularly for those in the higher prevalence categories. Often these high reported figures are due to some confounding factor (e.g., focusing on athlete populations known for a prevalence of doping), or limited to a specific substance in a specific population (e.g., anabolic steroids), or in time. Nonetheless, the overall picture from Table 2 suggest that evidence for doping prevalence is most robust in the low end with the majority of the included studies showing doping prevalence below 5% in both SAP and SRP for current and recent use. For self-reported lifetime use, the majority of the evidence still falls in the low (0%-5%) range but multiple studies were also found in the higher prevalence ranges.

Study Group	**Timeframe**	**Doping prevalence range**	**Count of studies in range**	**References**
SAP	Not specified except in Petróczi [80]	0% - 5%	12	25, 112, 113, 115, 117-119, 121, 123-126
		5% - 10%	4	25, 120, 124, 125
		10% - 20%	4	80, 116, 122, 124,
		20% - 30%	1	124
		30% - 40%	1	124
		40% - 50%	1	124
		> 50%	-	-
SRP	Within the past 12 months	0% - 5%	15	27, 50, 52, 54, 58, 72-73, 79, 96, 97, 101, 103, 105, 106
		5% - 10%	6	47, 54, 82, 105, 108, 109
		10% - 15%	3	104, 105, 108
		15% - 20%	4	55, 75, 105, 108
		20% - 30%	4	48, 55, 105, 108
		30% - 40%	4	34, 39, 55, 105
		> 50%	2	55, 110
	All (including current and past 12 months)	0% - 5%	40	23, 27, 30, 35, 36, 38, 40, 43, 44, 46, 49, 50, 51, 54, 56, 59, 61, 63, 66, 69-74, 77, 84, 88, 91, 96-98, 100, 101, 103, 105, 106, 108, 111, 114
		5% - 10%	18	27, 29, 32, 33, 47, 53, 54, 60, 62, 66, 69, 79, 87, 88, 90, 105, 108, 109
		10% - 15%	12	24, 28, 42, 66-68, 78, 80, 93, 104, 105, 108
		15% - 20%	8	24, 41, 42, 55, 75, 82, 105, 108

		20% - 30%	11	24, 48, 55, 64, 82, 83, 89, 99, 105, 107, 108
		30% - 40%	6	34, 55, 65, 83, 99, 107
		40% - 50%	5	55, 83, 99, 107, 110
		> 50%	7	39, 81, 83, 85, 92, 95, 110

Table 2*. Strength of evidence for ranges of doping prevalence (studies with sample ranges across prevalence rate categories appear multiple times). Self-reported prevalence (SRP); Sample analysis for prevalence (SAP).*

3.3. Population Represented in Studies with Evidence for Doping Prevalence

Of the 105 studies, 94 studies reported the size of the population surveyed, which totaled 102,515 and ranged from 8 to 13,914. An additional 14 studies only reported the number of samples analyzed but not the number of unique athletes providing samples, which totaled 1,484,554 samples and ranged from 42 to 1,347,214. Two studies did not report the total athletes or samples included in the population, though the data was used in the study findings.

3.4. Doping Prevalence by Gender

Of the 105 studies, 85 reported the gender[3] of the athletes included in the prevalence data. This meant that of the 102,515 participants, gender was reported for 81,041 athletes (79%). The reported gender of athletes identified 59,015 men (72.8%) and 22,026 women (27.2%). With the sample analysis, of the 1,484,544 samples, only 57,956 (4%) of the samples reported gender while no gender information was provided for 96% of the samples analyzed. The reported gender of samples provided by athletes identified 42,442 from men (73.2%) and 15,514 from women (26.8%). Only three studies reported athletes not identifying as either man or woman.

[3] Gender is the term used in official documents and reporting throughout sport governing bodies such as the International Olympic Committee, the Court of Arbitration for Sport, and the World Anti-Doping Agency to classify competition categories for men and women. As this evidence synthesis only related to competitive sport, the manuscript reflects the categorizations used by the competitive sport governing bodies.

For evidence of doping prevalence by gender, authors concluded that no meaningful synthesis could be drawn at the data reported in the surveyed studies. Table 3 offers a qualitative synthesis of doping prevalence by gender, but it must be interpreted with caution due to the limited number of studies reporting doping prevalence by gender.

Gender	Prevalence range	Count of studies	References
Male	<5%	14	27, 34, 38, 44, 49, 59, 66, 70, 79, 96, 98, 105, 114, 115
	5%-10%	11	26, 27, 53, 66, 70, 79, 80, 87, 90, 105, 109
	10%-20%	8	24, 66, 102, 104, 105, 108, 116, 122
	20%-30%	6	24, 55, 64, 80, 99, 105
	30%-40%	4	39, 55, 99, 105
	40%-50%	2	55, 99
	> 50%	5	39, 55, 81, 95, 99
Female	< 5%	13	26, 27, 38, 46, 59, 66, 90, 96, 98, 105, 108, 111, 115
	5%-10%	8	55, 66, 70, 80, 100, 104, 105, 109
	10%-20%	6	55, 24, 80, 102, 105, 122
	20%-30%	4	55, 24, 105, 122
	30%-40%	1	55
	40%-50%	1	55
	> 50%	1	55

Table 3. *Level of evidence for doping prevalence by gender (timeframe: lifetime use, including last 12 months and current*

3.5. Doping Prevalence by Sport

Establishing doping prevalence for specific sports would offer important insight into doping behavior. However, the authors concluded that no meaningful data could be generated by examining which sports were included in the studies. Indeed, there often appeared to be no methodological considerations related to sports participations. Many studies did not list the sports that athletes played, others gave several examples of a class but did not list all of the sports (e.g. “Team sports: Football, Basketball, etc.”), and few provided the number of athletes in their study that played a particular sport. Studies often remained unclear how much a particular sport (e.g. “cycling”) was actually sampled since studies would not indicate how many athletes

from the sport participated in the study. Therefore, the existing research offers little evidence that can help depict the doping behavior of athletes in a particular sport.

3.6. Doping Prevalence by Country

Establishing doping prevalence for specific countries should prove very useful for anti-doping efforts. The studies identified 34 countries while 15 studies involved an international mix of athletes. Studies that reported athletes from more than one country were classified as international. The largest surveyed country was the United States (with 26 studies) while 20 countries had only one study (see Table 4 for detailed description).

Country	Number	References
United States	26	27, 34, 35, 38, 39, 41, 43, 44, 46, 47, 50, 53, 56, 59-61, 73, 79, 84, 87, 90, 91, 95, 98, 99, 106
International	12	85, 101, 102, 110, 112, 113, 121, 122, 123, 124, 125, 126
Germany	8	42, 88, 96, 103, 104, 107, 108, 109
Greece	6	30-32, 62, 63, 75
Iran	5	57, 58, 67, 68, 86
United Kingdom	5	28, 33, 55, 97, 127
Nigeria	4	1, 4, 63, 72
Spain	3	70, 83, 111
Australia	2	51, 54
Brazil	2	40, 119
Canada	2	45, 48
Hungary	2	80, 93
Romania	2	74, 94
Saudi Arabia	2	25, 114
South Africa	2	37, 49
Sweden	2	54-55
Belgium	1	120
Bosnia and Herzegovina	1	81
Cameroon	1	26
Croatia	1	100
Czech Republic	1	71
Denmark	1	105
France	1	66
Guadeloupe	1	116

India	1	52
Italy	1	117
Jordan	1	89
Kenya	1	76
Macedonia	1	82
Malaysia	1	36
Mexico	1	118
Norway	1	115
Not Reported	1	29
Sri Lanka	1	92
Turkey	1	78
Uganda	1	72
TOTAL	105	

Table 4. *Number of studies with doping prevalence data by country.*

Determining which countries have a higher doping prevalence can assist in prevention and detection efforts. However, authors concluded that no meaningful synthesis could be drawn at the country level from the surveyed studies. Table 5 offers a qualitative synthesis at a regional level, but it must be interpreted with caution due to the limited number of studies and differing evidence for doping prevalence across studies. Among the 12 studies with international samples only one offers country-level breakdown [102] for Australia, US and UK. These were included in each relevant prevalence range. Studies were omitted from the table if they did not offer country-level breakdown [29, 101, 110, 112, 113, 121-123, 125, 126, 110,] or where the doping prevalence for the countries are not identified [124].

Region	Prevalence range	Count of studies	References
Africa	0% - 5%	3	49, 72, 77
	5%-10%	2	26, 69
	10%-20%	1	24
	20%-30%	1	24
	30%-40%	1	23
	40%-50%	-	-
	> 50%	-	-
Asia	0% - 5%	6	25, 36, 52, 57, 58, 114
	5%-10%	1	25

	10%-20%	1	67, 68 (same data)
	20%-30%	1	89
	30%-40%	-	-
	40%-50%	-	-
	> 50%	1	92
Australia	0% - 5%	2	51, 54
	5%-10%	1	54
	10%-20%	1	102
	20%-30%	-	-
	30%-40%	-	-
	40%-50%	-	-
	> 50%	-	-
Europe	0% - 5%	14	30, 63, 66, 70, 71, 74, 88, 94, 96, 100, 103, 111, 115, 117
	5%-10%	7	32, 33, 62, 66, 88, 109, 120
	10%-20%	11	28, 42, 55, 75, 78, 80, 82, 93, 102, 104, 108
	20%-30%	6	42, 55, 64, 105, 107, 108
	30%-40%	4	55, 65, 107, 108
	40%-50%	1	55
	> 50%	3	55, 83, 81
North America	0% - 5%	14	27, 35, 38, 43, 44, 46, 50, 56, 61, 73, 84, 91, 98, 106
	5%-10%	7	27, 47, 53, 60, 79, 87, 90
	10%-20%	3	27, 41, 102
	20%-30%	2	27, 48
	30%-40%	3	27, 34, 99
	40%-50%	1	99
	> 50%	3	39, 95, 99
South America & Caribbean	0% - 5%	3	40, 118, 119
	5%-10%	-	-
	10%-20%	1	116
	20%-30%	-	-
	30%-40%	-	-
	40%-50%	-	-
	> 50%	-	-

Table 5: *Level of evidence for doping prevalence by geographical region (Timeframe: lifetime use, including last 12 months and current; Turkey is included in Europe; Mexico is included in South America & Caribbean)*

Furthermore, a number of studies did not include how many athletes from a specific nation were included in their study, which prevented a weighted analysis that accounts for the differences between larger and smaller studies.

Of the studies reporting regional indicators, several nations produced far more prevalence studies than other nations while the vast majority of nations had no specific studies of athletes in their region. For example, the United States had 26 studies, while the next closest nations were Germany with eight studies and Greece with six studies. Although 11 studies involved surveys of international athletes (e.g. analyses of WADA laboratory statistics), of the 206 countries with a national Olympic committee, 172 had no specific studies involving their athletes. This indicates that even if an evidence synthesis of regions was possible, significant gaps exist for many geographic regions. Furthermore, the skewed distribution shown in Table 2 is likely due to the location of the researchers rather than the doping issue in the country *per se*.

3.7. Doping Prevalence by Level of Sport

Understanding doping prevalence for specific levels of sport also could provide important insights into doping behavior. Authors found that the studies did not present any consistent manner to divide athletes. While some studies focused on elite international athletes, others examined youth athletes (ages under 18). Others looked at amateur competitive athletes (e.g. people entering a triathlon that are not professional triathletes) or competed for their university team. However, upon further analysis, it became clear that many of the populations overlapped. For example, a number of youth athletes were also elite international athletes while college

athletes ranged from regional to international levels of their sport. With too many differences between studies, authors determined that no meaningful mapping could indicate doping prevalence for different levels of sport.

Additionally, authors discovered that the doping prevalence for para-athletes provided another methodologically difficult item to report. Despite para-athletes' sustained presence throughout sport, few studies specifically identified para-athletes as participants. Thus, authors were unable to confidently assert when para-athletes had been included. Several studies also likely included para-athletes in the data set, such as those using sample analyses from anti-doping laboratories, but did not note para-athlete inclusion. This presents a complicated, if not absent, picture for doping in para-sports. While para-athletes appear to be under studied relative to their presence in sport, the existing literature does not report sufficient information to determine whether this population is being accurately included or to estimate the doping prevalence in para-sports. Further complication arose that studies report prevalence of specific practices such as "boosting" (the practice of triggering autonomic dysreflexia via self-inflicted pain in athletes with spinal cord injuries at T6 or above) as doping when in fact "boosting" is not a prohibited method by WADA' List of Prohibited Substances and Methods.

3.8. Survey Questions to Establish Use

Of the 88 studies using survey questions, 51 included the question used to establish doping use (see Supplementary Material Appendix S3). However, the questions varied significantly across studies thus no further analysis could reveal trends or continuity across studies. Some studies only specifically referenced anabolic steroids or a limited number of prohibited substances. Other studies used broad terms such as "doping" or "prohibited performance enhancing

substances." Some made efforts to differentiate substances permitted for therapeutic use while others did not discriminate between prohibited and permitted practices.

3.9. Timeframe of Use

The time frame for doping provides important information for prevalence studies. Of the 88 studies using surveys, only 19 provided clear time frames (e.g. last 12 months). Other studies used terms such as "currently using" or "ever used," which provide some notion of time frame but may not accurately capture the nature of doping behavior (Supplementary Material Appendix S3). Thirty-two studies using surveys did not indicate a timeframe regarding use. The varied reporting and standardization for timeframe of use prevented any further analysis.

3.10. Studies with Evidence of Doping Prevalence Over Time

The evidence synthesis did not restrict studies to a particular date range. The earliest identified published study appeared in 1975 and this review included papers published up to and covering 2019. Of the 105 studies, 4 were published between 1975 and 1989. For five-year periods between 1990 and 2019, the number of studies increased every period, with 35 studies conducted between 2015-2019 (see Figure 2).

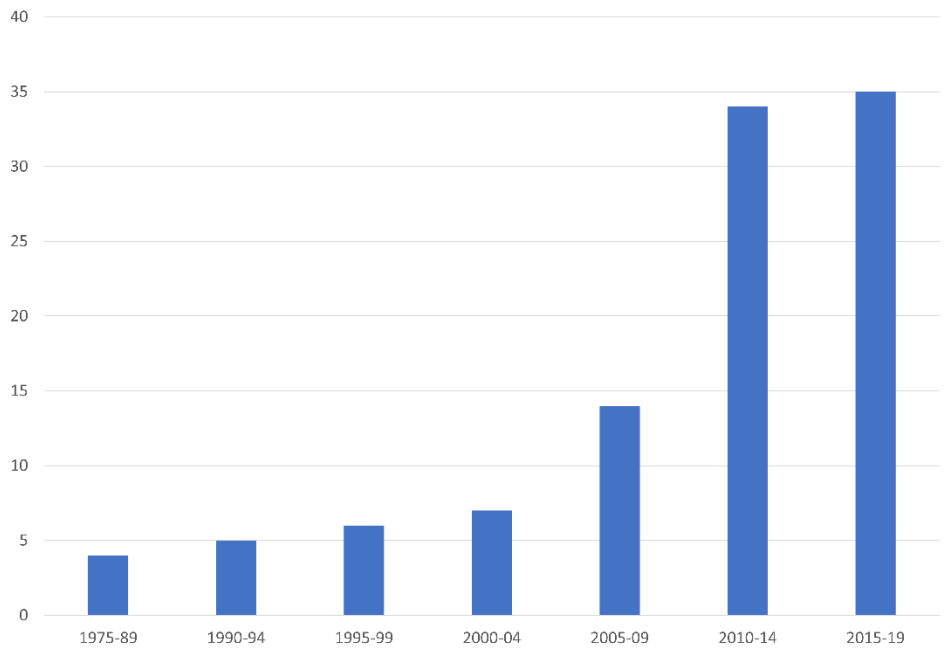

Figure 2. *Number of studies with data for doping prevalence grouped in 5-year periods from 1975-1989 to 2015-2019.*

The decade between 2010 to 2019 produced a total of 69 of the 105 studies ranging between 3 to 12 studies with an average of 6.9 studies per year. Though early, the trend suggests that after a period of initial growth (which is expected from the similar trend in doping research in general) the annual number of studies may have started to plateau.

3.11. Quality Assessment

The results of the quality assessment indicated 20 studies rated "High," 60 rated "Moderate," and 25 rated "Low" for their evidence of doping prevalence (Table 4). The complete scoring for each study is included in Supplementary Material Appendix S4. The five highest rated studies [110, 103, 34, 42, 30] used either the Random Response Technique (RT) or its variants; or the Direct Survey (DS) methods to determine doping prevalence. The RT, which is specifically designed for prevalence studies on sensitive topics, had the highest percentage of "High" ratings (See discussion in 4.2.3 below).

Study Group	Study Method	Studies (N)	Quality		
			High	Moderate	Low
SRP	DS	79	16%	59%	24%
SAP	TF	10	10%	50%	40%
SRP	RT	10	40%	50%	10%
SAP	AD	3	33%	33%	33%
SAP	BP	3	33%	67%	0%
SRP	NS	1	0%	100%	0%
SRP	QI	1	0%	100%	0%
SAP	HS	1	0%	100%	0%

Table 6. *Quality assessment by study design. Self-reported prevalence (SRP); Sample analysis for prevalence (SAP); Direct survey (DS); Random response technique (RT); Testing figure (TF); Athletes biological passport (BP); Anti-doping rule violation (AD); Qualitative interview (QI); Hair sample analysis (HS); Network scale-up (NS).*

The quality assessment also indicated that the quality of evidence for doping prevalence is increasing along with its quantity (See Figure 3).

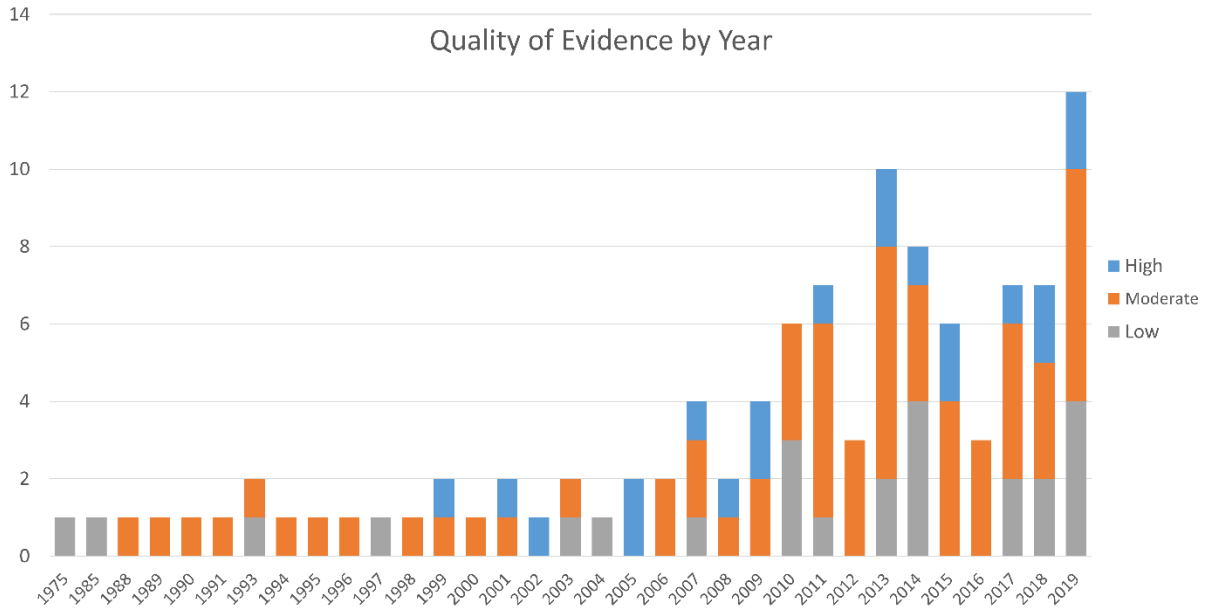

Figure 3. *The Quality of Evidence for Doping Prevalence by Publication Year.*

The QUAD-SRP and QUAD-SAP also produced evidence that indicates the quality of reporting on doping prevalence. With the quality assessment, it is important to keep in mind that quality of the data available for doping prevalence in these studies was assessed, not the quality of the study. As it was noted earlier, the majority of the studies included was not set out for establishing prevalence. Rather, information on doping use in the investigated cohort of athletes was included as part of the sample characteristics for the study. When reviewing how each question affected the QUAD-SRP rating, the item analysis showed Question #5 and #10 had the greatest predictive value for "High" compared to "Low" studies. Too few studies were included in the QUAD-SAP to perform an item analysis. Quality of reporting within the studies are reported as proportions for the percent of scores for each question on the QUAD-SRP for the 89 survey studies (Table 7) and the QUAD-SAP for the 17 sample analysis studies (Table 8).

Question #	QUAD-SRP Criteria	Yes	No	Partial	NA
1	Addresses the extent to which the findings from the study can be generalized to the population from which the study subjects are derived.	52.8%	47.2%	0.0%	0.0%
2	Addresses the extent to which survey used validated instruments standardized to the field.	47.2%	24.7%	28.1%	0.0%
3	Addresses the extent to which the sample size is appropriate for establishing doping prevalence.	59.6%	5.6%	34.8%	0.0%
4	Considers data quality in terms of whether method is appropriate to obtain evidence for doping behavior.	92.1%	5.6%	2.2%	0.0%
5	Assesses whether the measurement of doping was objective.	100.0%	0.0%	0.0%	0.0%
6	Assess whether participants understood what was meant by doping when answering questions.	43.8%	56.2%	0.0%	0.0%
7	Assess whether participants understood the time frame for doping behavior being measured.	64.0%	36.0%	0.0%	0.0%
8	Considers clarity of instructions and assurance for non-exposure (for indirect estimation methods only).	3.4%	9.0%	0.0%	87.6%
9	Addresses bias in the measurement of the outcomes in a study.	77.8%	22.2%	0.0%	0.0%
10	Addresses bias in the interpretation of ambiguous substances (e.g., cannabis, alcohol, prescription drugs).	28.1%	66.3%	0.0%	5.6%
11	Addresses potential bias in data collection procedure through loss of control.	60.7%	36.0%	0.0%	3.4%
12	Addresses whether studies have applied adjustment for confounding in the analysis.	16.9%	22.5%	0.0%	60.7%
13	Considers whether the dataset was altered retrospectively (i.e., altered after data collection completed, during data analysis).	20.2%	4.5%	0.0%	75.3%
14	Considers whether participants were sufficiently protected during data collection.	78.7%	19.1%	0.0%	2.2%
15	Assesses whether the information provided in the paper is sufficient to allow a reader to make an unbiased assessment of the findings of the study.	31.5%	53.9%	14.6%	0.0%
16	Assess whether the study controlled for non-compliant responses.	16.9%	83.1%	0.0%	0.0%
17	Assess whether study reports accurate evidence for doping behavior.	87.6%	12.4%	0.0%	0.0%

Table 7*. Quality Analysis of Doping of Doping evidence-Self-Reported Prevalence (QUAD-SRP) with extraction results reported in the appropriate columns as percentages. Instructions and criteria for scoring of studies are included in Supplementary Material Appendix S1.*

Question #	QUAD-SAP Criteria	Yes	No	Partial
1	Addresses the extent to which the findings from the study can be generalized to the population from which the study subjects are derived.	100%	0%	0%
2	Addresses the extent to which sample analysis can indicate population.	35%	35%	29%
3	Addresses the extent to which the sample size is appropriate for establishing doping prevalence.	35%	6%	59%
4	Addresses the extent to which the reported findings are appropriate for establishing doping prevalence.	76%	0%	24%
5	Considers data quality in terms of whether method is appropriate to obtain evidence for doping behavior.	41%	59%	0%
6	Assess whether participants were competitive athletes.	100%	0%	0%
7	Assess whether the time frame for doping behavior was measured.	88%	12%	0%
8	Considers whether the dataset included samples that were not doping behaviors.	35%	65%	0%
9	Considers the extent of inadvertent doping.	0%	100%	0%
10	Considers the confounding factors when prevalence of doping is inferred for population level.	18%	82%	0%

Table 8*. Quality Analysis of Doping of Doping evidence-Sample Analysis for Prevalence (QUAD-SAP) with extraction results reported in the appropriate columns as percentages. Instructions and criteria for scoring of studies are included in Supplementary Material Appendix S2.*

3.12 Doping prevalence for studies assessed as “High” for quality of evidence

Of the 20 studies evaluated by either the QUAD-SAP or QUAD-SRP as “High” for quality of evidence, 10 of the studies reported doping prevalence between 0% and 5%. However, it is difficult to conclude whether this is evidence of a trend. The study method may have influenced the reported doping prevalence as 8 of the 10 studies in the lowest range (0-5% range) used a Direct Survey method while only 3 studies using a Direct Survey method reported estimates

above 5%. At the same time, the study with the highest quality assessment score for evidence [110]used a "randomized response technique"[4] and reported the highest doping prevalence range of the group, with doping prevalence estimated between 43.6-57.1%. Readers are cautioned against applying the 'higher must be closer to the truth' criterion without due consideration of the method, population as well as quality and generalizability of the evidence. Equally, the studies included in the "High" synthesis for doping prevalence evidence involved narrow populations, such as adolescent athletes in the United States [43] or European male football players [121], reported prevalence for only one substance such as anabolic androgenic steroid use among high school American football players [87] or among pre-adolescent athletes [89]. Thus, additional factors, such as methodological bias, differences in sample populations, or differences in substances measured suggest researchers should be cautious about viewing the number of studies reporting doping prevalence between 0-5% as the best representation of doping prevalence in competitive sport.

Doping prevalence range	**Count of studies in range**	**References**
0% - 5%	10	35, 43, 66, 72, 73, 96, 98, 103, 106, 121
5% - 10%	5	29, 30, 66, 87, 109
10% - 15%	2	78, 104
15% - 20%	1	122
20% - 25%	0	-
25% - 30%	0	-
30% - 35%	1	34
35% - 40%	0	-
40% - 45%	1	110
45% - 50%	1	110
> 50%	1	110

[4] A multitude of indirect estimation models exists. In the applied literature, these are often referred to as 'randomized response technique', even though not all models rely on randomization. For simplicity and to avoid confusion, we accepted this terminology for the review but noting its inaccuracy.

Table 9. *Prevalence ranges for studies with quality assessment rated prevalence evidence "High"*

4. DISCUSSION

Reported doping prevalence rates in competitive sport rates ranged between 0% and 73%, with most falling under 5%. In total, the included studies involved 102,515 competitive athletes (72.8% men and 27.2% women) in 35 countries, but competitive level varied from interschool and club-level to international. The evidence synthesis detailed above marks a milestone in the doping prevalence literature. For the first time, an expert group has reviewed an exhaustive collection of studies reporting doping prevalence for competitive sport, synthesized the evidence, and provided quality assessment. Admittedly, the disparate evidence does not provide the desired clarity on the past or current prevalence of doping. Even among the studies rated to be "High" quality, the diverse methods, terms, populations, date ranges, and limitations undermine the confidence in the aggregation. However, in discussing the limitations of this evidence synthesis, a clearer understanding of the pathway forward emerges that can assist researchers to produce better data on doping prevalence.

4.1. Scope of Doping Prevalence Estimates

The differences between study designs and survey questions make comparing or synthesizing doping prevalence figures difficult. For example, studies that only examined anabolic steroid use in the last three months cannot be synthesized with studies that surveyed all prohibited performance-enhancing substance over an athlete's career. Additionally, some studies included recreational drug use alongside prohibited performance-enhancing substance use or only examined one prohibited substance such as Anabolic Androgenic Steroids as evidence of doping.

Finally, the evidence synthesis included many studies that did not report establishing a doping prevalence as an intended aim of the study. This likely affected the data collected as well as the manner the data was recorded and reported. These issues will be discussed more in the recommendations for best practices (Section 5).

4.2. Review of Methods for Establishing Doping Behavior

As mentioned above, most study designs do not specifically seek to determine doping prevalence yet generate evidence of doping behavior for a specific population. As such, they provide (or at least could provide) sample estimates of doping prevalence. Review of the studies indicated four different methods that can indicate doping prevalence and had been used in more than one study. Since repeated use indicates some measure of adoption by the field, we noted instances where only one study used a specific method. Of the 108 methods used, 89 (82.4%) involved surveys, of which 79 (73.1%) used some form of direct questions and 10 (9.3%) used some variant of the randomized response techniques. Yet three studies used multiple measures to examine doping behavior, which can improve the quality of evidence by better triangulating and informing doping prevalence estimates. As reporting methods become standardized, scholars can increasingly rely on a growing body of data to compare the reliability of methods across populations.

4.2.1. Sample Analysis

Sample analysis typically involved determining doping prevalence by directly screening athletes' samples for indications of doping. In addition to the doping control data examined above, sample analysis includes studies that analyze samples for prohibited substances and its metabolites or indication of doping through changes in values across multiple samples. Current methods for

sample analysis involve samples of an athlete's urine, blood, or hair as well as Athlete Biological Passports. The evidence synthesis identified 10 (9.3%) studies using sample analysis.

Sample analysis has several strengths. Scientifically valid tests can screen for all known prohibited substances and for markers of prohibited methods, samples from athletes with TUEs can be identified so as to be separated from doping behaviors, and samples avoid problems of false reporting from athletes that either lie or become confused during an interview or survey. The evidence synthesis identified nine studies that used sample analysis to provide some evidence of doping behavior.

Sample analysis is easy for researchers to use thanks to the standardized reporting requirements developed by WADA for its various anti-doping organizations, yet it also has challenges. Establishing the prevalence for a population can be difficult because the number of individuals in the sample may not be reported. Athletes often give multiple samples and not every athlete provides the same number of samples while some may provide no samples. For example, the United States Anti-Doping Agency reported for 2017 that it analyzed 9,820 total samples, but it only tested 3,576 athletes. Of the athletes tested, 1,741 athletes were only tested once, while three athletes were tested 16 times [128].

Furthermore, studies using sample analysis must distinguish between Adverse Analytical Findings (AAF) and an actual Anti-Doping Rule Violation (ADRV). AAFs are simply a positive test for a prohibited substance reported by the accredited anti-doping laboratories, which means that athletes with therapeutic use exemptions do not receive an ADRV. For example, USADA had 136 AAFs (1.4% of the samples analyzed) in 2017, but only 67 became ADRVs (0.7% of the samples analyzed). But for prevalence purposes, the best indication would be ADRVs (67) of the

tested population (3576), which is a prevalence of 1.8% [128]. The added challenge here is separating ADRVs for prohibited substance or method use from other forms of rule violations such as tampering, refusing to give samples, evading testing, failing the whereabouts requirement, trafficking and assisting doping.

A more significant challenge for sample analysis comes from athletes' efforts to avoid detection in doping controls as well as the general limitations that affect the doping control system. Much evidence has indicated that doping athletes take steps to avoid providing a positive sample or to alter a potentially positive sample [129]. Furthermore, most sample analyses draw upon laboratory testing, which is a tool designed for anti-doping rule enforcement and not specifically intended for prevalence. This means data drawn from anti-doping testing may overrepresent samples from suspicious athletes targeted for testing or the samples required from athletes finishing in the top three of a competition and is unlikely to accurately represent the broader athlete population.

For these reasons, most of the studies relying on sample analysis scored lower for their quality of evidence for doping prevalence. One exception to this is a recent paper [122], which presents a novel approach specifically designed to use ABP data for doping prevalence estimation. Thus, only three studies in the SAP group achieved a "High" rating. Establishing the number of athletes represented in the population and avoiding conflating AAFs with doping behavior are vital steps before researchers should use sample analysis to estimate doping prevalence.

4.2.2. Direct Self-reports in Surveys

Direct self-report surveys involve an athlete reporting their own doping behavior in response to a direct survey question. Of the studies that reported their survey methods, 79 (73.1%) used a

direct questionnaire that anonymously asked the athlete to indicate any doping behavior. Few of these studies indicated efforts to ensure truthful responses from participants beyond offering anonymity thus leaving some doubt as to whether athletes answered the questions honestly and accurately. Even on anonymous questionnaires, athletes may have some concern about honestly reporting doping behavior while having little incentive to do so. Concerns that someone might see their answers or that somehow their answers may be linked back to them create an incentive not to report doping behavior. This means direct questionnaires may underestimate doping behavior and potentially lead to lower doping prevalence estimates.

Even if researchers assume athletes responded truthfully, variations in study designs prevent data synthesis. The scope and timeframe of reporting doping differ across studies. Timeframe ranges from "current," "last season," "last 12 months," up to lifetime with "ever" or "in the past" (e.g. Kabiri et al [58], Pitsch [106], Gallucci et al. [47], Yesalis [99]). Responses also ranged from binary response ("Yes/No" in Kisaalita and Robinson [60]) to extent ("I do not use doping, / I use doping from time to time / I use doping on a regular basis" in Rodek et al [81]) or variations of a Likert scale ("0=No, I don't use/5=Yes, I usually use" in Hejabi et al. [52]). Given the nature of doping practices among competitive sport, extent or Likert scale questions appear poorly suited to determine doping prevalence while qualifying time frame seems particularly useful.

4.2.3. Surveys Utilizing Indirect Estimation Models

Athletes are surveyed using a type of indirect estimation survey, commonly referred to as randomized response technique (RT). The search identified 10 studies (9.3%) using various RT surveys to establish doping behavior for a population (See Table 1). Though there are several ways such RT surveys are conducted, the common feature is the added statistical "noise" to the

survey response which makes linking affirmative answer to the doping question impossible. This statistical "noise" can only be considered at the sample level. This approach offers protection over and above anonymity for athletes because only they know the full picture. It also provides a relatively inexpensive and quick survey that can include large numbers of athletes.

RTs also have limits and drawbacks. They can be complicated to design correctly and their complexity may confuse athletes. The evidence synthesis also identified a range of statistical methods and interpretations such as "cheater detection" (Pitsch [106]) such that the prevalence estimates across studies cannot always be synthesized. The varying ways of reporting the results make this unfamiliar way of measuring prevalence figures even more difficult to fathom, especially for non-mathematicians. Finally, RTs require large populations of athletes to provide a statistically meaningful prevalence estimate. RTs are unlikely to yield meaningful evidence with small samples making the approach unsuitable for some studies. It is because the relatively large sample is needed to obtain the expected distribution of the added "noise" and distinguish the prevalence estimates from this "noise." For example, if the "noise" is one's birthday we know that half of the population has the birthday in the first six months [130] but a sufficiently large sample is needed to obtain this distribution in the investigated sample. As a rule of thumb, generally samples of 500 or larger are used for sufficient power to detect small but meaningful prevalence rates [131].

Additionally, RT addresses one element of socially desirable responding, which is the fear of exposure. It does not, however, motivate athletes to report their behavior honestly. As there is no omnipotent metric to compare the answers against, some degree of uncertainty will persist regarding athletes' truthfulness in answering the questions. Still, as has been shown in various topics [132], RT studies mitigate limitations found in other social scientific efforts so may

provide more reliable measurements. Additionally, research cautions against automatically assuming that a higher estimate is closer to the "true" prevalence because lying about a sensitive issue such as doping depress prevalence figures obtained through direct questioning more than the RT prevalence estimate. Although RT estimates tend to be higher than direct questioning, this assumption, referred to as 'the higher is better' rule, may not always the case [133].

4.2.4. Qualitative Interviews

One study administered a survey face-to-face and used semi-structured interviews to discuss doping behaviors for specific sub-populations [111]. Since each qualitative interview requires significant time, it involved smaller sample size of 8 interview participants. The study also reported 0% doping prevalence. This points to potential limits inherent to qualitative interviews. First, studies with a small sample size can be more susceptible to statistical variation. Second, the intimate nature of face-to-face interview may lead some participants to lie about their doping behavior in order to avoid admitting to the interviewer their own participation in a socially stigmatized behavior. Thus qualitative interviews present significant methodological limits as a tool for establishing doping prevalence.

4.2.5. Anti-Doping Rule Violations

Three studies employed anti-doping rule violations (ADRVs) as a method to examine doping behavior for a population of competitive athletes [125-127]. Examining ADRVs to determine doping prevalence offers several advantages. ADRVs are always public and often easily accessed through WADA, international sport federations and national anti-doping organizations. They also avoid some problems identified with the sample analysis because they identify individuals rather than positive samples, which may over represent doping athletes.

However, ADRVs have limits as evidence of doping prevalence. ADRVs have the same basic problem associated with sample analysis, which stem from the challenge to identify doping behavior designed to avoid detection as an ADRV. Additionally, ADRVs reflect the result of a quasi-legal administrative process shaped by the resources to prosecute and defend the charges of an anti-doping rule violation. Issues ranging from unintentional ingestion of prohibited substances through contaminated supplements, failure to obtain a valid TUE, mistaken use of prohibited medications as well as failing whereabouts requirements, trafficking and abetting illustrate that not all ADRVs qualify as doping behavior, whereby athletes intentionally use prohibited substances to enhance performances in competitive sports. Moreover, the financial and legal challenges in defending against an ADRV should caution researchers not to treat an ADRV as unqualified proof of doping.

Finally, researchers may also struggle to determine the population size from which the ADRVs emerged to establish doping prevalence. For example, Aubel et al. [125] calculated the total number of ADRVs for professional cycling between 2005 to 2016 but did not determine the total number of professional cyclists for that period, which is necessary to establish the doping prevalence. In many cases, the number of athletes registered to an international sport federation or a national anti-doping organization may not be available to researchers and even with the information, researchers may struggle to determine with consistency the number of athletes actually controlled for ADRVs. The variations in testing regimes across international sports federations and national anti-doping organizations and the use of targeted testing may further limit the usefulness of ADRVs. Furthermore, not all ADRVs necessarily involve athletes intentionally using a prohibited substance to enhance performance. Some ADRVs involve

coaches while others include failed whereabouts reporting or failure to follow TUE policy, which may not indicate intentional doping behavior.

4.2.6. Network Scale Up

Currently, only one study on doping used the network scale up method (NS) (See Table 1). This method is typically used in hard to reach populations such as with HIV positive people or sex workers [15]. Surveyors ask individuals to estimate how many people they know and then how many of those known people do they know are doing the specific behavior. Combining these estimates with the known population allows researchers to mathematically model the prevalence of the particular behavior.

Since doping athletes can be considered a hard to reach population, this method may provide researchers another way to estimate doping prevalence with several advantages. Researchers may better estimate doping prevalence for an entire population without having to survey large numbers of people as required in the direct survey or random response techniques. However, NS also has limits. If doping behavior is completely hidden, athletes may not actually know about their teammates' or competitors' behavior thus leading to underestimation of doping behavior. Therefore, as with ADRVs, the NS requires further development from researchers to determine if the method is viable for doping prevalence studies.

4.3. Strength of Evidence

The increase in studies reporting doping behaviors in competitive sport indicates a growing interest and improved understanding generated by researchers that may ultimately lead to better quality evidence of doping prevalence. The earliest published study that achieved a "High" rating was published in 1999. However, in the 19 years that followed, only 19 studies achieved a

similar mark. That amounts to one “High” rating of evidence for doping prevalence a year. Such a rate is low for a complicated, fluid, and hidden practice that continues to change over time. Significant steps must be taken by authors and editors to improve the quality of evidence.

5. “BEST PRACTICE” RECOMMENDATIONS

The preceding evidence synthesis shows that estimates of doping prevalence in competitive sports have only modestly illuminated pockets of the hidden practices. In part, researchers have limited the field’s impact because a number of studies asked the wrong questions, used inappropriate research designs or methods, or reported incomplete or inadequate information on their study. Such issues reduce the quality of evidence, limits the value of the data, and contributes to “avoidable waste” in research production [134]. More advanced fields in the medical sciences have developed standardized methods for data production and dissemination that serve as a useful model for preventing waste while improving the quality of evidence [135].

Informed by the results of the evidence synthesis and quality assessment, the authors who were members of WADA’s Working Group on Doping Prevalence, in consultation with senior members of WADA, identified a set of guidelines and best practice recommendations that will standardize and improve the quality of the evidence generated and disseminated so as to better support the scholarly community and policymakers wishing to use their research. Collaboration between academic researchers and policymakers is identified as a best-practice for evidence synthesis [14]. However, these best practice recommendations are intended to support scholarly community’s research aims to produce better quality of evidence indicating doping prevalence rather than to address any policymaker’s agenda.

5.1 Recommendations for Research Design

Most studies providing prevalence data are often not intended to be strictly prevalence studies. However, when researchers establish how many athletes in a defined group are doping, they have the opportunity to contribute prevalence data. The following guidelines should assist researchers studying doping practices in competitive sports to produce high-quality studies while contributing to data to better illustrate doping prevalence.

1. *Studies should clearly differentiate "doping" from behaviors that do not involve athletes intentionally using prohibited substances to enhance performances in competitive sports, such as using nutritional supplements, therapeutic medications, and illegal or recreational drug use.*

WADA's *Prohibited List* provides a comprehensive and understandable list of prohibited doping substances and methods. However, some researchers still conflate using prohibited doping substances for performance-enhancement with non-prohibited therapeutic medications, illegal drugs for recreation, and nutritional supplements. Examples identified in the evidence synthesis included studies identifying morphine, methadone, opium, phenobarbital, and barbiturates alongside anabolic steroids; confusing anabolic steroids that are prohibited at all times with glucocorticosteroids that are not prohibited out of competition; or psychoactive drug use in general as "performance enhancing drugs" used by the participants (see Ajayi-Vincent and Olanipekun [24] and Pereira and Sardela [119]) without any discussion of possible legitimate therapeutic purposes or the fact that some of these are not even prohibited performance-enhancing substances. Thus, a claimed prevalence may not accurately depict doping prevalence for their surveyed population since it includes activities not considered doping. The same can be said for a number of sample analysis studies that employ adverse analytical findings (AAF)

and/or atypical findings rather than anti-doping rule violations (ADRV). As discussed above, AAFs represent presence of a prohibited substance or method in athletes' samples and are not systematically considered anti-doping rule violations and thus should not be considered "doping." However, WADA and anti-doping organizations can significantly assist study authors using sample analysis by providing reports that indicate (and retroactively update) the number of AAFs from samples that become ADRVs. In all reporting, study authors should make efforts to report athletes engaging in doping practices separate from athletes using a substance for therapeutic treatment or for recreational purposes to assist in identifying doping prevalence rates.

2. *When possible, authors should provide a direct estimate for doping behavior within a specified sample and use keywords "doping prevalence" in publications to identify the data*.

The evidence synthesis revealed that much doping prevalence data is going unidentified while some prevalence is even going unreported. Despite a thorough keyword search, the vast majority of the data emerged through snowball sampling and the research team's knowledge. The gap stemmed largely from surveys that indicated doping prevalence but did not identify the figure as such. In other studies, the research team clearly gathered the doping prevalence for their participants but did not actually report the number (e.g. Soltanabadi et al. [86] and Whitaker and Backhouse [127]). While it was clear the researchers have the information for prevalence, the data was omitted from the manuscript and thus research could not contribute to an understanding of doping prevalence. Study authors can also assist in disseminating better quality of evidence by reporting and referencing the percentage of athletes in a study that dope as "doping prevalence," which will help researchers identify and use their study's findings. Study authors should better appreciate the value of prevalence figures to other researchers and policy makers.

3. *Surveys of athletes' doping behavior should provide a defined frame of reference for any doping practices*

The evidence synthesis identified that surveys of athletes' doping behavior varied widely in the timeframe for the activity. While some surveys asked athletes, "Have you ever doped?" others provided more helpful questions such as, "Have you doped in the last 12 months" or "Have you doped in the last season?", or asking about current use "Do you currently use prohibited substances to enhance your sport performance"? Particularly problematic questions attempted to retrofit Likert-type scale measurements of doping (e.g. whether an athlete doped "a large amount," "a moderate amount," or "not at all") which were poor indicators of doping behavior. Given the unique nature of doping practices, an athlete may have doped once to enhance performance early in their career but not have doped in the years that followed. The fluid nature of doping behavior and anti-doping interventions mean that a defined time frame will provide researchers with higher quality evidence. For time-frame, we recommend using "last 30 days," "last 12 months," and/or "ever" for lifetime use, unless a precise timeframe is required for addressing a specific research question.

4. *Authors providing indirect estimates for doping behavior through proxy methods should avoid referring to data as "doping prevalence."*

Frequently used proxy indicators for doping behavior, albeit excluded from this evidence synthesis, include intention to dope, doping susceptibility or willingness. Research indicates these constructs have links to actual doping behavior [12] and such questions can provide useful insights. However, study authors should avoid classifying such responses as doping prevalence since they do not provide a prevalence figure. The same applies to response-time based implicit

measures (e.g., Autobiographical Implicit Association Test [136], which are pursued as a measure free of socially desirable responding. Overwhelming evidence indicates that implicit estimates are poor indicators of actual behavior by members of the group [9]. Given the difference between what doping prevalence attempts measure and what indirect proxy methods actually measure, the term doping prevalence would inaccurately represent data gathered through indirect estimates.

If researchers wish to establish doping prevalence, survey questions should ask about the behavior (i.e. "Use(d) prohibited performance enhancing substances and/or methods without Therapeutic Use Exemption") and should not be exchanged synonymously with related but distinct social cognitive measures such as consideration, willingness, likelihood or intention. Equally, if the prevalence question uses the phrase "doping," researchers should define for the participants what constitute "doping."

5. *Projected prevalence estimates should not be interpreted and reported as prevalence.*

Indirectly estimated doping behavior via projective questions should not be confused with doping prevalence. With projective questions, researchers may ask athletes, "What percentage of your opponents do you think doped in the last 12 months?". Assuming that respondents do not have the accurate information, their responses to this question is a guess that is heavily influenced by an egocentric bias. Projected prevalence of a behavior is on one hand influenced by the respondents' environment and beliefs; and on the other hand, by the behavior in question [9]. Undesirable behavior that shared with others tends to be overestimated in a phenomenon known as the "false consensus effect" [137], whereas shared proportion of desirable behavior is typically underestimated, in a phenomenon known as "uniqueness bias" [138]. Either way, these

estimations are more revealing about the person making the estimates than the actual population prevalence, and influenced by the relative closeness of the estimation (e.g., guessing about their own teammates, own sport, own country or the competitors locally or globally). In other studies, athletes were asked if they personally knew athletes who dope. This, again, is revealing about the athletes' environment and the perception of doping use but the number of athletes who report knowing someone is not evidence for doping prevalence, especially not in a small and defined sample where it is likely that multiple athletes 'know' the same doper. Researchers wish to extrapolate from the number of dopers known to population prevalence are advised to use established methodology such as the Network Scale-Up technique [15].

6. *Studies using randomized/fuzzy response technique in surveys should take noncompliance into account.*

A high rate of noncompliance in survey data derived from randomized/fuzzy response techniques (e.g., Crosswise Model, Forced Response Technique, Randomized Response Technique, Single Sample Count, and Unrelated Question Model) has been documented in the literature on doping prevalence and beyond. More often than not, in studies that estimate the proportion of the sample that is noncompliant, it is assumed that noncompliance is deliberate, motivated, labelled as "cheating" and pool admitted behavior and noncompliance together. Yet noncompliance can also be caused by the complexity of the survey technique where respondents do not understand the instructions or do not make the effort to read the instructions carefully. This means that only a proportion of the noncompliant responses are deliberate lies about the sensitive behavior in question (e.g., doping). Therefore, unless there is evidence for the source of noncompliance, survey results should be reported as proportion of positive cases (i.e., admitted doping use), proportion of negative cases (declared no use) and proportion of noncompliant responses. When

noncompliance is considered to adjust the estimation of the behavior of interest, both unadjusted and adjusted estimation should be reported.

7. *Studies of doping prevalence should gather and report the level of competition and national identity for athletes surveyed.*

While the WADA code differentiates between "International," "National," and "Non-National" level athletes, additional distinctions among competition levels likely provide researchers with important prevalence information. For example, useful levels may include specifying the inclusion of para-athletes, age-group athletes at both the youth and senior or "master's" level, and club, recreational, or amateur level athletes. Such distinctions can better support determining prevalence for specific populations. However, dividing athletes into levels proves methodologically challenging and practically difficult for some studies where the distinction is not obvious. In certain cases, an athlete may qualify both as an age group athlete and as a national or international level athlete. To address this issue, researchers may wish to provide multiple metrics when working with athletes that represent more than one specific level. Researchers should always include a description of the distance (if present) in performance level relative to the international elite level in the studies sport(s), which constitutes the highest possible level, of the studied population. Likewise, a clear description of the country or countries represented by the athletes in the study. For multiple countries, it should include the number of athletes representing each country.

8. *Studies should attempt to report gender and consider gender representation in studies.*

The evidence synthesis indicated research included more male athletes (73%) than female athletes (27%). Depending on the situation, a gender may be overrepresented in a research study.

However, researchers should consider whether the gender representation in the study provides an accurate reflection of sport participation and make efforts to appropriately sample the gender represented in sport participation. Reporting should also indicate the number of men and women represented in the study as well as gender-nonconforming athletes when appropriate. As doping prevalence may be different by gender, studies with mixed gender may wish to report prevalence by gender identity. Study authors are encouraged to provide information about gender to help prevent underrepresentation of a gender both in research studies and while compiling evidence for doping prevalence. When possible, authors should present prevalence statistics for men and women as most competitive sports treat these as separate populations. However, authors must balance recording and reporting of gender information with any promise of confidentiality or anonymity in data reporting. If reporting of gender data threatens to reveal participants identity, then study authors may omit such reporting. For studies employing sample analysis to determine prevalence, WADA and anti-doping organizations can significantly assist study authors by providing the percent of samples drawn from each gender in the compilation of laboratory reports.

9. *Studies focusing on one or several specific sports should identify the sports being surveyed in line with the sports/discipline classification used by sport governing bodies such as International Olympic Committee or WADA classification.*

While some studies may survey all sports, such as those using WADA's compiled laboratory statistics, many other studies included athletes from a limited number of sports. Reporting the participants specific sports helps to determine the amount of data for a particular sport. The evidence synthesis indicated some sports, such as weightlifting and cycling, are vastly

overrepresented in the prevalence literature while little research has reported doping prevalence for many other sports leaving large gaps in the literature.

Following standardized sport reporting can also prevent confusion about which sport was actually surveyed (e.g. “hockey” could be either “field hockey” or “ice hockey”). Also, some sports such as biathlon are separate from skiing, while researchers in skiing may wish to designate specific disciplines such as Alpine, Cross-Country, or Ski Jumping. Authors should also note para-sports separately when athletes compete separately. For example, a survey of both tennis and wheelchair tennis players should be listed as a study of “tennis” players, ideally reporting sport disciplines both separately and in total. Finally, authors must balance recording and reporting of sport and discipline with any promise of confidentiality or anonymity in data reporting. If reporting of sport or discipline data threatens to reveal participants identity, then study authors may report information in ways that ensure participants remain anonymous or omit reporting of information that compromises anonymity, but if possible, keep the information on record to make it available upon request for meta-analyses.

10. *When using sample analysis to establish prevalence, studies should distinguish between the number of tests and the number of individuals tested.*

The evidence synthesis demonstrated that all but one study relying upon sample analysis failed to differentiate between the number of tests and the number of athletes tested. As previously discussed, athletes often provide more than one sample. For example, an anti-doping organization may not test all of their athletes the same number of times; some athletes may only be tested once while others may provide multiple tests. Such practices mean that surveying the results of 10,000 tests is not the same as surveying 10,000 athletes. For this reason, the evidence

synthesis separated the prevalence reporting for sample analysis from athlete surveys. While study authors should report prevalence for the number of athletes, WADA can significantly facilitate this reporting by having its national anti-doping organization provide the same formation (e.g. number of total tests and total number of athletes tested) in their annual reports to WADA to be included in the Laboratory Reports. When working with historical data, it is recommended that adverse analytical findings and atypical findings are triangulated with ADRVs to avoid inflation in prevalence owing to contamination or therapeutic uses.

5.2. Reporting Guidelines

Authors (and editors) seeking to publish research on doping behavior in competitive sport should adhere to guidelines for ethical reporting of data such as those provided by the Vancouver Convention for Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals [139]. In particular, study authors should avoid reproducing data or fragmenting data across multiple publications, as the practice risks distorting the scientific literature, especially in reviews or meta-analyses. When data is reproduced for appropriate reasons, authors should follow Vancouver Convention guidelines for citing the original dataset.

In addition, authors should follow these reporting guidelines specific to research on doping behavior in order to effectively communicate higher quality research. Even studies not specifically focused on doping prevalence should be expected to follow these guidelines, as the guidelines will increase the study's impact and relevance in a growing body of research.

Studies should include the following information:

- Number of athletes surveyed or tested

- Number of athletes identifying by gender (Men/Women) and prevalence rates by gender, if appropriate.
- Timeframe when the data was conducted
- Clear operational definition of 'doping behavior' used in the study with indication of how respondents were informed of this definition
- Method used to determine doping behavior (e.g., sample analysis, direct survey, indirect survey)
- Timeframe considered for doping behavior (e.g., current, last 3 months, last 12 months, career)
- Number and percent of athletes indicated as doping during specified timeframe
- Sports represented in survey/testing, corresponding to classification used by international sporting federations and para sporting federations
- Level of athletes surveyed and/or tested, corresponding to classifications of athlete pathways (e.g. International, National, Talented, Youth, etc.) used by international sporting organizations.
- Nations represented by athletes surveyed and/or tested
- If data is part of larger data set or previously published, authors should cite original source for data.
- If prevalence is estimated using randomized/fuzzy response techniques in survey or estimated from data from the Athlete Biological Passport, report confidence/credible interval or standard error of measurement; and identify clearly which one is reported.

These reporting guidelines will be further enhanced by following the "best practice" recommendations as detailed above. Combined, the best practice recommendations and the

reporting guidelines should not only improve the quality and usefulness of doping research but also allow for more useful meta-analyses and evidence surveys that better reveal the prevalence of doping behavior.

6. CONCLUSION

While researchers have advanced the understanding of doping prevalence, especially since 2010, the field still has a long way until it can begin producing high quality doping prevalence estimates. The challenges to producing such high-quality research are surmountable if the field of doping research matures and coordinates as a scientific community. Such coordination is vital. The actual prevalence of doping will never be a question answered by one research team using one methodology. Indeed, it will always require geographically diverse research teams and necessitate multiple methods. However, if all parties interested in determining doping prevalence in competitive sport commit to developing and standardizing best practices and reporting guidelines, then better estimates of doping prevalence will more clearly illuminate the presently opaque practice of doping in competitive sport.

ACKNOWLEDGEMENT

The work was conducted as part of the Working Group on Doping Prevalence of the World Anti-Doping Agency (WADA). The authors thank Tony Cunningham, Marcia MacDonald, and Olivier Rabin for their critical review and constructive comments on the manuscript; and Annie Bachman for her assistance in extracting and compiling data.

COMPLIANCE WITH ETHICAL STANDARDS

Funding

No sources of funding were used to assist in the preparation of this article.

Conflicts of interest

This paper represents part of the work by the World Anti-Doping Agency Working Group on Doping Prevalence conducted between September 2017 and December 2019, but WADA had no control over the drafting or content of this manuscript. John Gleaves, Andrea Petróczi, Olivier De Hon, Martial Saugy and Maarten Cruyff served as members of the Working Group (2017-2019) and they prepared this paper in their capacity as Working Group members, in collaboration with DF and EM. The Working Group members receive no salary for their work but expenses related to the travel for work were covered. Andrea Petróczi received grant funding from WADA previously as part of the Social Science Research Program, served as a member of the first Working Group on Doping Prevalence (2011-2012); and is currently involved in providing analysis and evaluation support for WADA's Outreach Program in an unpaid advisory role. Martial Saugy worked at the Swiss Laboratory for Doping analyses (LAD, Centre Hospitalier Universitaire Vaudois (CHUV), Lausanne, Switzerland) until 2016 and received funding from WADA Science Department prior to his involvement in this project. Olivier De Hon works for the National Anti-Doping Authority Netherlands. Dirk Folkerts and Emmanuel Macedo declare they have no conflicts of interest relevant to the content of this review.

Consent to participate

Not applicable

Availability of data and materials

The definitions, questions, and rater criteria for the Quality Assessment of Doping evidence – Self-Reported Prevalence (QUAD-SRP) and the Quality Assessment of Doping evidence – Sample Analysis of Prevalence (QUAD-SRP) are available in Supplementary Material Appendix S1 and S2, respectively. All extracted data from the studies is available in Supplementary Material Appendix S3. The complete scoring for all studies is available in Supplementary Material Appendix S4. All other datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authorship contributions

Andrea Petróczi served as senior author on the project, conceptualized the study, led the development of quality assessment criteria, contributed to collating and synthesizing the independent quality assessments, contributed to the literature search, supervised Dirk Folkerts and contributed to drafting the manuscript. John Gleaves drafted the manuscript, contributed to the development of quality assessment criteria, contributed to collating and synthesizing the independent quality assessments as well as the literature search and supervised Emmanuel Macedo. Folkerts conducted the initial literature search, contributed to developing the quality assessment criteria and conducted independent quality assessment for all included studies under the supervision of Petróczi. Olivier De Hon conducted independent quality assessment, contributed to developing the quality assessment criteria and literature search. Macedo conducted independent quality assessment under the supervision of Gleaves. The best practice recommendations were formulated by Petróczi, Gleaves, De Hon, Martial Saugy, and Maarten Cruyff. All authors read and critically commented on the manuscript and approved the final version of the manuscript.

REFERENCES

1. De Hon O, Kuipers H, van Bottenburg M. Prevalence of doping use in elite sports: a review of numbers and methods. Sports Med. 2015;45(1):57-69.

2. Maennig W. Inefficiency of the anti-doping system: cost reduction proposals. Subst Use Misuse. 2014;49(9):1201-5.

3. Grohmann K. WADA to ask commercial sponsors for money in anti-doping fight. Reuters. 2019 5 November.

4. Pielke Jr R. Gather data to reveal true extent of doping in sport: drug cheats will not be tackled properly until anti-doping agencies do more to assess the scale of the problem scientifically. Nature. 2015;517(7536):529-30.

5. Pielke R. Assessing doping prevalence is possible: So what are we waiting for? Sports Med. 2018;48(1):207-9.

6. Ohl F, Fincoeur B, Lentillon-Kaestner V, Defrance J, Brissonneau C. The socialization of young cyclists and the culture of doping. International Review for the Sociology of Sport. 2015;50(7):865-82.

7. Curtis A, Gerrard D, Burt P, Osborne H. Drug misuse in sport: A New Zealand perspective. NZ Med J. 2015;128:62.

8. Dimeo P, Taylor J. Monitoring drug use in sport: The contrast between official statistics and other evidence. Drugs: education, prevention and policy. 2013;20(1):40-7.

9. Petróczi A, Backhouse SH, Barkoukis V, Brand R, Elbe A-M, Lazuras L, et al. A matter of mind-set in the interpretation of forensic application. International Journal of Drug Policy. 2015;26(11):1142-3.

10. Blank C, Kopp M, Niedermeier M, Schnitzer M, Schobersberger W. Predictors of doping intentions, susceptibility, and behaviour of elite athletes: a meta-analytic review. SpringerPlus. 2016;5(1):1333.

11. Backhouse S, Whitaker L, Patterson L, Erickson K, McKenna J. Social psychology of doping in sport: a mixed studies narrative synthesis. Project report. World Anti-Doping Agency, Montreal Canada. 2016.

12. Ntoumanis N, Ng JY, Barkoukis V, Backhouse S. Personal and psychosocial predictors of doping use in physical activity settings: a meta-analysis. Sports Med. 2014;44(11):1603-24.

13. Sutherland WJ, Wordley CFR. A fresh approach to evidence synthesis. Nature. 2018;558(7710):364-6. doi:10.1038/d41586-018-05472-8.

14. Donnelly CA, Boyd I, Campbell P, Craig C, Vallance P, Walport M, et al. Four principles to make evidence synthesis more useful for policy. Nature. 2018;558(7710):361-4. doi:10.1038/d41586-018-05414-4.

15. Bernard HR, Hallett T, Iovita A, Johnsen EC, Lyerla R, McCarty C, et al. Counting hard-to-count populations: the network scale-up method for public health. Sexually Transmitted Infections. 2010;86(Suppl 2):ii11-ii5.

16. Monin B, Norton MI. Perceptions of a fluid consensus: Uniqueness bias, false consensus, false polarization, and pluralistic ignorance in a water conservation crisis. Personality and Social Psychology Bulletin. 2003;29(5):559-67.

17. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 2009;6(7):e1000097.

18. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. Systematic Reviews. 2016;5(1):210.

19. Sagoe D, Molde H, Andreassen CS, Torsheim T, Pallesen S. The global epidemiology of anabolic-androgenic steroid use: a meta-analysis and meta-regression analysis. Ann Epidemiol. 2014;24(5):383-98.

20. Greenhalgh T, Peacock R. Effectiveness and efficiency of search methods in systematic reviews of complex evidence: audit of primary sources. BMJ. 2005;331(7524):1064-5.

21. Hoy D, Brooks P, Woolf A, Blyth F, March L, Bain C, et al. Assessing risk of bias in prevalence studies: modification of an existing tool and evidence of interrater agreement. J Clin Epidemiol. 2012;65(9):934-9.

22. Wong WC, Cheung CS, Hart GJ. Development of a quality assessment tool for systematic reviews of observational studies (QATSO) of HIV prevalence in men having sex with men and associated risk behaviours. Emerging themes in epidemiology. 2008;5(1):23.

23. Afolayan J, Adegboyega J. Knowledge and use of performance enhancing drugs among Nigeria elite athletes. IOSR J Appl Chem. 2012;1(5):31e8.

24. Ajayi-Vincent O, Olanipekun J. An assessment of prohibited methods of doping among athletes in tertiary institutions in southwest, Nigeria. Kuwait chapter of Arabian journal of business and management review. 2013;33(858):1-6.

25. Al Ghobain M, Konbaz MS, Almassad A, Alsultan A, Al Shubaili M, AlShabanh O. Prevalence, knowledge and attitude of prohibited substances use (doping) among Saudi sport players. Subst Abuse Treat Prev Policy. 2016;11:14. doi:10.1186/s13011-016-0058-1.

26. Ama P, Betnga B, Moor VA, Kamga J. Football and doping: study of African amateur footballers. Br J Sports Med. 2003;37(4):307-10.

27. Anderson WA, Albrecht RR, McKeag DB, Hough DO, McGrew CA. A national survey of alcohol and drug use by college athletes. The Physician and Sportsmedicine. 2016;19(2):91-104. doi:10.1080/00913847.1991.11702148.

28. Backhouse SH, Whitaker L, Petroczi A. Gateway to doping? Supplement use in the context of preferred competitive situations, doping attitude, beliefs, and norms. Scand J Med Sci Sports. 2013;23(2):244-52. doi:10.1111/j.1600-0838.2011.01374.x.

29. Barkoukis V, Lazuras L, Tsorbatzoudis H, Rodafinos A. Motivational and sportspersonship profiles of elite athletes in relation to doping behavior. Psychol Sport Exerc. 2011;12(3):205-12. doi:10.1016/j.psychsport.2010.10.003.

30. Barkoukis V, Lazuras L, Lucidi F, Tsorbatzoudis H. Nutritional supplement and doping use in sport: possible underlying social cognitive processes. Scand J Med Sci Sports. 2015;25(6):e582-8. doi:10.1111/sms.12377.

31. Barkoukis V, Lazuras L, Ourda D, Tsorbatzoudis H. Are nutritional supplements a gateway to doping use in competitive team sports? The roles of achievement goals and motivational regulations. J Sci Med Sport. 2020;23(6):625-32. doi:10.1016/j.jsams.2019.12.021.

32. Barkoukis V, Lazuras L, Tsorbatzoudis H, Rodafinos A. Motivational and social cognitive predictors of doping intentions in elite sports: an integrated approach. Scand J Med Sci Sports. 2013;23(5):e330-40. doi:10.1111/sms.12068.

33. Boardley ID, Smith AL, Mills JP, Grix J, Wynne C. Empathic and self-regulatory processes governing doping behavior. Front Psychol. 2017;8:1495. doi:10.3389/fpsyg.2017.01495.

34. Buckman JF, Farris SG, Yusko DA. A national study of substance use behaviors among NCAA male athletes who use banned performance enhancing substances. Drug Alcohol Depend. 2013;131(1-2):50-5. doi:10.1016/j.drugalcdep.2013.04.023.

35. Buckman JF, Yusko DA, White HR, Pandina RJ. Risk profile of male college athletes who use performance-enhancing substances. J Stud Alcohol Drugs. 2009;70(6):919-23. doi:10.15288/jsad.2009.70.919.

36. Ming Chiang L, Shamsuddin AF, Tuan Mahmood TM. Knowledge, attitude and practice on doping of Malaysian student athletes. Asian Journal of Pharmaceutical and Clinical Research. 2018;11(5). doi:10.22159/ajpcr.2018.v11i5.23598.

37. Coopoo Y, Jakoet I. Substance abuse and knowledge thereof among elite South African sportspeople. South African Journal of Sports Medicine. 2000;7(3):10-3.

38. Corbin CB, Feyrer-Melk SA, Phelps C, Lewis L. Anabolic steroids: a study of high school athletes. Pediatr Exerc Sci. 1994;6(2):149-58.

39. Curry LA, Wagman DF. Qualitative description of the prevalence and use of anabolic androgenic steroids by United States powerlifters. Percept Mot Skills. 1999;88(1):224-33.

40. Silva PRPd, Maranhao Neto GA, Figueiredo VC, Santos AMPVd, Jacob MHVM, Rose EHd, et al. Doping survey in the youth school games in Brazil. Revista Brasileira de Medicina do Esporte. 2017;23(6):436-40. doi:10.1590/1517-869220172306163303.

41. Dezelsky T, Toohey J, Shaw R. Non-medical drug use behaviour at five United States universities: a 15-year study. Bull Narc. 1985;37(2-3):49-53.

42. Dietz P, Ulrich R, Niess A, Best R, Simon P, Striegel H. Prediction profiles for nutritional supplement use among young German elite athletes. Int J Sport Nutr Exerc Metab. 2014;24(6):623-31. doi:10.1123/ijsnem.2014-0009.

43. Dodge T, Jaccard JJ. Is abstinence an alternative? Predicting adolescent athletes' intentions to use performance enhancing substances. J Health Psychol. 2008;13(5):703-11. doi:10.1177/1359105307082460.

44. Dodge T, Stock M, Litt D. Judgments about illegal performance-enhancing substances: reasoned, reactive, or both? J Health Psychol. 2013;18(7):962-71. doi:10.1177/1359105312445079.

45. Donahue EG, Miquelon P, Valois P, Goulet C, Buist A, Vallerand RJ. A motivational model of performance-enhancing substance use in elite athletes. Journal of Sport and Exercise Psychology. 2006;28(4):511-20.

46. Elliot DL, Goldberg L, Moe EL, DeFrancesco CA, Durham MB, Hix-Small H. Preventing substance use and disordered eating: initial outcomes of the ATHENA (athletes targeting healthy exercise and nutrition alternatives) program. Arch Pediatr Adolesc Med. 2004;158(11):1043-9.

47. Gallucci AR, Martin RJ. Misuse of prescription stimulant medication in a sample of college students: examining differences between varsity athletes and non-athletes. Addict Behav. 2015;51:44-50. doi:10.1016/j.addbeh.2015.07.004.

48. Goulet C, Valois P, Buist A, Côté M. Predictors of the use of performance-enhancing substances by young athletes. Clin J Sport Med. 2010;20(4):243-8.

49. Gradidge P, Coopoo Y, Constantinou D. Prevalence of performance-enhancing substance use by Johannesburg male adolescents involved in competitive high school sports. Archives of Exercise in Health and Disease. 2011;2(2):114-9. doi:10.5628/aehd.v2i2.102.

50. Green GA, Uryasz FD, Petr TA, Bray CD. NCAA study of substance use and abuse habits of college student-athletes. Clin J Sport Med. 2001;11(1):51-6.

51. Gucciardi DF, Jalleh G, Donovan RJ. An examination of the sport drug control model with elite Australian athletes. J Sci Med Sport. 2011;14(6):469-76. doi:10.1016/j.jsams.2011.03.009.

52. Hejabi A, Manouchehri J, Tojari F. Determining validity and reliability of doping behavior measurement instrument in young athletes cociety. J Appl Environ Biol Sci. 2015;5(3):246-53.

53. Horn S, Gregory P, Guskiewicz KM. Self-reported anabolic-androgenic steroids use and musculoskeletal injuries: findings from the Center for the Study of Retired Athletes health survey of retired NFL players. Am J Phys Med Rehabil. 2009;88(3):192-200. doi:10.1097/PHM.0b013e318198b622.

54. Jalleh G, Donovan RJ, Jobling I. Predicting attitude towards performance enhancing substance use: a comprehensive test of the sport drug control model with elite Australian athletes. J Sci Med Sport. 2014;17(6):574-9.

55. James RA, Nepusz T, Naughton DP, Petróczi A. A potential inflating effect in estimation models: cautionary evidence from comparing performance enhancing drug and herbal hormonal supplement use estimates. Psychol Sport Exerc. 2013;14(1):84-96.

56. Judge LW, Bellar D, Petersen J, Lutz R, Gilreath E, Simon L, et al. The attitudes and perceptions of adolescent track and field athletes toward PED use. Performance Enhancement & Health. 2012;1(2):75-82. doi:10.1016/j.peh.2012.04.002.

57. Kabiri S, Cochran JK, Severson R, Shadmanfaat SM, Rahmati MM, Sharepour M. Social and personal controls and performance enhancing drug use: toward an explanation of doping activity among professional athletes in Rasht, Iran. Deviant Behavior. 2018;39(11):1483-96D. doi:10.1080/01639625.2018.1486057.

58. Kabiri S, Shadmanfaat SM, Donner CM. Using control balance theory to examine sports doping among professional athletes in Iran. Journal of Drug Issues. 2019;49(3):493-511. doi:10.1177/0022042619837757.

59. Kersey RD. Anabolic-androgenic steroid use among California community college student-athletes. Journal of Athletic Training. 1996;31(3):237.

60. Kisaalita NR, Robinson ME. Attitudes and motivations of competitive cyclists regarding use of banned and legal performance enhancers. J Sports Sci Med. 2014;13(1):44.

61. Krowchuk DP, Anglin TM, Goodfellow DB, Stancin T, Williams P, Zimet GD. High school athletes and the use of ergogenic aids. Am J Dis Child. 1989;143(4):486-9.

62. Lazuras L, Barkoukis V, Rodafinos A, Tzorbatzoudis H. Predictors of doping intentions in elite-level athletes: a social cognition approach. Journal of Sport and Exercise Psychology. 2010;32(5):694-710.

63. Lazuras L, Barkoukis V, Tsorbatzoudis H. Toward an integrative model of doping use: an empirical study with adolescent athletes. J Sport Exerc Psychol. 2015;37(1):37-50. doi:10.1123/jsep.2013-0232.

64. Lindqvist A-S, Moberg T, Eriksson BO, Ehrnborg C, Rosén T, Fahlke C. A retrospective 30-year follow-up study of former Swedish-elite male athletes in power sports with a past anabolic androgenic steroids use: a focus on mental health. Br J Sports Med. 2013;47(15):965-9.

65. Ljungqvist A. The use of anabolic steroids in top Swedish athletes. Br J Sports Med. 1975;9(2):82.

66. Lorente FO, Peretti-Watel P, Grelot L. Cannabis use to enhance sportive and non-sportive performances among French sport students. Addict Behav. 2005;30(7):1382-91. doi:10.1016/j.addbeh.2005.01.019.

67. Manouchehri J, Tojari F. Examining the conceptual model: relationships between sport orientation, doping attitude and doping behavior in Iranian elite martial artists. Eur J Exp Biol. 2013;3(2):175-82.

68. Manouchehril J, Tojari F. Development and validation of instruments to measure doping attitudes and doping beliefs. Eur J Exp Biol. 2013;3(2):183-6.

69. Molobe ID. Knowledge, attitude and practice on drug abuse among sports men and women in Lagos State, Nigeria. International journal of medicine and medical sciences. 2012;2(3):77-85.

70. Morente-Sanchez J, Zandonai T, Zabala Diaz M. Attitudes, beliefs and knowledge related to doping in different categories of football players. J Sci Med Sport. 2019;22(9):981-6. doi:10.1016/j.jsams.2019.05.010.

71. Mudrak J, Slepicka P, Slepickova I. Sport motivation and doping in adolescent athletes. PLoS One. 2018;13(10):e0205222. doi:10.1371/journal.pone.0205222.

72. Muwonge H, Zavuga R, Kabenge PA. Doping knowledge, attitudes, and practices of Ugandan athletes': a cross-sectional study. Subst Abuse Treat Prev Policy. 2015;10:37. doi:10.1186/s13011-015-0033-2.

73. Naylor AH, Gardner D, Zaichkowsky L. Drug use patterns among high school athletes and nonathletes. Adolescence. 2001;36(144):627.

74. Nica-Badea D. Proximity Influence of Environment Social Support Specific Sportsmen and Risk Factors of Intention and Behavior Doping. Pensee. 2014;76(7).

75. Ntoumanis N, Barkoukis V, Gucciardi DF, Chan DKC. Linking coach interpersonal style with athlete doping intentions and doping use: a prospective study. J Sport Exerc Psychol. 2017;39(3):188-98. doi:10.1123/jsep.2016-0243.

76. Ogama D, Omariba E, Sakwa MM. Exploring the impact of knowledge and attitudes on doping behaviour among athletics youthful Kenyan long-distance runners. International Academic Journal of Law and Society. 2019;1(2):35-47.

77. Ohaeri JU, Ikpeme E, Ikwuagwu PU, Zamani A, Odejide OA. Use and awareness of effects of anabolic steroids and psychoactive substances among a cohort of Nigerian professional sports men and women. Human Psychopharmacology: Clinical and Experimental. 1993;8(6):429-32.

78. Özdemir L, Nur N, Bagcivan I, Bulut O, Sümer H, Tezeren G. Doping and performance enhancing drug use in athletes living in Sivas, mid-Anatolia: a brief report. J Sports Sci Med. 2005;4(3):248.

79. Petróczi A, Uvacsek M, Nepusz T, Deshmukh N, Shah I, Aidman EV, et al. Incongruence in doping related attitudes, beliefs and opinions in the context of discordant behavioural data: in which measure do we trust? PLoS One. 2011;6(4):e18804.

80. Petroczi A. Attitudes and doping: a structural equation analysis of the relationship between athletes' attitudes, sport orientation and doping behaviour. Subst Abuse Treat Prev Policy. 2007;2:34. doi:10.1186/1747-597X-2-34.

81. Rodek J, Sekulic D, Pasalic E. Can we consider religiousness as a protective factor against doping behavior in sport? J Relig Health. 2009;48(4):445-53. doi:10.1007/s10943-008-9207-9.

82. Ruzdija K, Gontarev S, Lazar T. Doping attitudes and doping behavior in Macedonian professional wrestler. Research in Physical Education, Sport, and Health 2018;7(1):55-60.

83. Sanchez-Oliver AJ, Grimaldi-Puyana M, Dominguez R. Evaluation and behavior of Spanish bodybuilders: doping and sports supplements. Biomolecules. 2019;9(4). doi:10.3390/biom9040122.

84. Schneider D, Morris J. College athletes and drug testing: attitudes and behaviors by gender and sport. Journal of Athletic Training. 1993;28(2):146.

85. Silvester J. Anabolic steroids at the 1972 Olympics! Coach and Athletic Director. 2006;76(3):8.

86. Soltanabadi S, Tojari F, Esmaeili MR. Variance analysis of sport motivation, doping attitude and behavior among pro athletes of team sports. Journal of Applied Environmental and Biological Sciences. 2015;5(3):254-58.

87. Stilger VG, Yesalis CE. Anabolic-androgenic steroid use among high school football players. J Community Health. 1999;24(2):131-45.

88. Striegel H, Ulrich R, Simon P. Randomized response estimates for doping and illicit drug use in elite athletes. Drug Alcohol Depend. 2010;106(2-3):230-2. doi:10.1016/j.drugalcdep.2009.07.026.

89. Tahtamouni LH, Mustafa NH, Alfaouri AA, Hassan IM, Abdalla MY, Yasin SR. Prevalence and risk factors for anabolic-androgenic steroid abuse among Jordanian collegiate students and athletes. Eur J Public Health. 2008;18(6):661-5. doi:10.1093/eurpub/ckn062.

90. Terney R, McLain LG. The use of anabolic steroids in high school students. Am J Dis Child. 1990;144(1):99-103.

91. Tricker R, Connolly D. Drugs and the college athlete: an analysis of the attitudes of student athletes at risk. J Drug Educ. 1997;27(2):105-19. doi:10.2190/E8U0-42UM-PBLK-96RH.

92. Uduwana I, Madushani A. Identifying the reasons for the misuse of ergogenic aid in national level athletes in Sri Lanka. European International Journal of Science and Humanities. 2014;1(2).

93. Uvacsek M, Nepusz T, Naughton DP, Mazanov J, Ranky MZ, Petroczi A. Self-admitted behavior and perceived use of performance-enhancing vs psychoactive drugs among competitive athletes. Scand J Med Sci Sports. 2011;21(2):224-34. doi:10.1111/j.1600-0838.2009.01041.x.

94. Vâjială G, Epuran M, Stanescu M, Potzaichin I, Berbecaru C. Relation between motivation and temptation for using the doping substances in high performance sports. Ovidius University Annals, Series Physical Education & Sport/Science, Movement & Health. 2010;10(2).

95. Wagman DF, Curry LA, Cook DL. An investigation into anabolic androgenic steroid use by elite US powerlifters. The Journal of Strength & Conditioning Research. 1995;9(3):149-54.

96. Wanjek B, Rosendahl J, Strauss B, Gabriel HH. Doping, drugs and drug abuse among adolescents in the state of Thuringia (Germany): prevalence, knowledge and attitudes. Int J Sports Med. 2007;28(4):346-53. doi:10.1055/s-2006-924353.

97. Whitaker L, Long J, Petroczi A, Backhouse SH. Using the prototype willingness model to predict doping in sport. Scand J Med Sci Sports. 2014;24(5):e398-405. doi:10.1111/sms.12148.

98. Wroble RR, Gray M, Rodrigo JA. Anabolic steroids and preadolescent athletes: prevalence, knowledge, and attitudes. Sport J. 2002;5(3):1-8.

99. Yesalis III CE, Herrick RT, Buckley WE, Friedl KE, Brannon D, Wright JE. Self-reported use of anabolic-androgenic steroids by elite power lifters. The Physician and Sportsmedicine. 1988;16(12):91-100.

100. Zenic N. Comparative analysis of substance use in ballet, dance sport, and synchronized swimming: results of a longitudinal study. Med Probl Perform Art. 2010;25(2):75-81.

101. Erickson K, Stanger N, Patterson L, Backhouse SH. Substance use in university sport: a cross-national study of student-athlete substance use behaviors and perceived responses to witnessing substance use. Performance Enhancement & Health. 2019;7(1-2). doi:10.1016/j.peh.2019.100151.

102. Boardley ID, Smith AL, Ntoumanis N, Gucciardi DF, Harris TS. Perceptions of coach doping confrontation efficacy and athlete susceptibility to intentional and inadvertent doping. Scand J Med Sci Sports. 2019;29(10):1647-54. doi:10.1111/sms.13489.

103. Dietz P, Ulrich R, Dalaker R, Striegel H, Franke AG, Lieb K, et al. Associations between physical and cognitive doping--a cross-sectional study in 2.997 triathletes. PLoS One. 2013;8(11):e78702. doi:10.1371/journal.pone.0078702.

104. Dietz P, Dalaker R, Letzel S, Ulrich R, Simon P. Analgesics use in competitive triathletes: its relationship to doping and on predicting its usage. J Sports Sci. 2016;34(20):1965-9. doi:10.1080/02640414.2016.1149214.

105. Elbe A-M, Pitsch W. Doping prevalence among Danish elite athletes. Performance Enhancement & Health. 2018;6(1):28-32.

106. Pitsch W, Emrich E. The frequency of doping in elite sport: results of a replication study. International Review for the Sociology of Sport. 2011;47(5):559-80. doi:10.1177/1012690211413969.

107. Pitsch W. Assessing and explaining the doping prevalence in cycling. Doping in Cycling: Interdisciplinary Perspectives. 2018.

108. Pitsch W, Emrich E, Klein M. Doping in elite sports in Germany: results of a www survey. European Journal for Sport and Society. 2007;4(2):89-102.

109. Seifarth S, Dietz P, Disch AC, Engelhardt M, Zwingenberger S. The prevalence of legal performance-enhancing substance use and potential cognitive and or physical doping in German recreational triathletes, assessed via the randomised response technique. Sports (Basel). 2019;7(12). doi:10.3390/sports7120241.

110. Ulrich R, Pope HG, Cléret L, Petróczi A, Nepusz T, Schaffer J, et al. Doping in two elite athletics competitions assessed by randomized-response surveys. Sports Med. 2018;48(1):211-9.

111. Morente-Sánchez J, Leruite M, Mateo-March M, Zabala M. Attitudes towards doping in Spanish competitive female Cyclists vs. Triathletes. Journal of Science and Cycling. 2013;2(2):40.

112. Aguilar M, Munoz-Guerra J, Plata MDM, Del Coso J. Thirteen years of the fight against doping in figures. Drug Test Anal. 2017;9(6):866-9. doi:10.1002/dta.2168.

113. Aguilar-Navarro M, Munoz-Guerra J, Del Mar Plara M, Del Coso J. Analysis of doping control test results in individual and team sports from 2003 to 2015. J Sport Health Sci. 2020;9(2):160-9. doi:10.1016/j.jshs.2019.07.005.

114. Al Ghobain M. The use of performance-enhancing substances (doping) by athletes in Saudi Arabia. J Family Community Med. 2017;24(3):151-5. doi:10.4103/jfcm.JFCM_122_16.

115. Bahr R, Tjørnhom M. Prevalence of doping in sports: doping control in Norway, 1977-1995. Clinical Journal of Sport Medicine. 1998;8(1):32-7.

116. Marchand A, Buisson C, Martin L, Martin JA, Molina A, Ressiot D. Report on an anti-doping operation in Guadeloupe: high number of positive cases and inferences about doping habits. Drug Test Anal. 2017;9(11-12):1753-61. doi:10.1002/dta.2185.

117. Mazzeo F, Monda M. Doping in Italy: an analysis of its spread in ten years. Biology and Medicine. 2016;08(01). doi:10.4172/0974-8369.1000263.

118. Mercado KS, Camacho EF, Rodríguez LB, Rodríguez MF, Mendoza NM, Velasco-Bejarano B. Banned substances and their incidence: a retrospective view of the national laboratory of prevention and doping control of Mexico. Adicciones. 2019;31(3):201-11.

119. Pereira HM, Sardela VF. Stimulant doping agents used in Brazil: prevalence, detectability, analytical implications, and challenges. Subst Use Misuse. 2014;49(9):1098-114. doi:10.3109/10826084.2014.907653.

120. Van Eenoo P, Delbeke F. The prevalence of doping in Flanders in comparison to the prevalence of doping in international sports. Int J Sports Med. 2003;24(08):565-70.

121. Vouillamoz M, Thom C, Grisdale R, Saugy M, Giraud S, Robinson N, et al. Anti-doping testing at the 2008 European football championship. Drug Test Anal. 2009;1(11-12):485-93. doi:10.1002/dta.105.

122. Faiss R, Saugy J, Zollinger A, Robinson N, Schuetz F, Saugy M, et al. Prevalence estimate of blood doping in elite track and field athletes during two major international events. Front Physiol. 2020;11:160.

123. Kuipers H, Moran J, Dubravcic-Simunjak S, Mitchell DW, Shobe J, Sakai H, et al. Hemoglobin level in elite speed skaters from 2000 up to 2005, and its relationship with competitive results. Int J Sports Med. 2007;28(1):16-20. doi:10.1055/s-2006-924055.

124. Sottas PE, Robinson N, Fischetto G, Dolle G, Alonso JM, Saugy M. Prevalence of blood doping in samples collected from elite track and field athletes. Clin Chem. 2011;57(5):762-9. doi:10.1373/clinchem.2010.156067.

125. Aubel O, Lefevre B, Le Goff JM, Taverna N. The team effect on doping in professional male road cycling (2005-2016). Scand J Med Sci Sports. 2019;29(4):615-22. doi:10.1111/sms.13384.

126. Maquirriain J. Epidemiological analysis of doping offences in the professional tennis circuit. J Occup Med Toxicol. 2010;5(1):30.

127. Whitaker L, Backhouse S. Doping in sport: an analysis of sanctioned UK rugby union players between 2009 and 2015. J Sports Sci. 2017;35(16):1607-13.

128. USADA. 2017 Athlete Test History. 2018. https://www.usada.org/news/athlete-test-history/. Accessed 21 April 2020.

129. Alquraini H, Auchus RJ. Strategies that athletes use to avoid detection of androgenic-anabolic steroid doping and sanctions. Mol Cell Endocrinol. 2018;464:28-33.

130. Petróczi A, Nepusz T, Cross P, Taft H, Shah S, Deshmukh N, et al. New non-randomised model to assess the prevalence of discriminating behaviour: a pilot study on mephedrone. Subst Abuse Treat Prev Policy. 2011;6(1):20.

131. Ulrich R, Schröter H, Striegel H, Simon P. Asking sensitive questions: a statistical power analysis of randomized response models. Psychol Methods. 2012;17(4):623.

132. Lensvelt-Mulders GJ, Hox JJ, Van der Heijden PG, Maas CJ. Meta-analysis of randomized response research: thirty-five years of validation. Sociological Methods & Research. 2005;33(3):319-48.

133. Höglinger M, Jann B. More is not always better: An experimental individual-level validation of the randomized response technique and the crosswise model. PLoS One. 2018;13(8).

134. Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. The Lancet. 2009;374(9683):86-9. doi:https://doi.org/10.1016/S0140-6736(09)60329-9.

135. von Niederhäusern B, Guyatt GH, Briel M, Pauli-Magnus C. Academic response to improving value and reducing waste: A comprehensive framework for INcreasing QUality In patient-oriented academic clinical REsearch (INQUIRE). PLoS Med. 2018;15(6):e1002580-e. doi:10.1371/journal.pmed.1002580.

136. Agosta S, Sartori G. The autobiographical IAT: a review. Front Psychol. 2013;4:519.

137. Mullen B, Atkins JL, Champion DS, Edwards C, Hardy D, Story JE, et al. The false consensus effect: a meta-analysis of 115 hypothesis tests. J Exp Soc Psychol. 1985;21(3):262-83.

138. Goethals GR, Messick DM, Allison ST. The uniqueness bias: studies of constructive social comparison. In: Suls J, Wills TH, editors, Social Comparison: Contemporary Theory and Research. Mahwah: Lawrence Erlbaum Associates; 1991. pp. 149–176.

139. International Committee of Medical Journal Editors. Recommendations for the conduct, reporting, editing, and publication of scholarly work in medical journals. 2018. http://www.icmje.org/icmje-recommendations.pdf. Accessed 21 April 2020.

COMPLIANCE WITH ETHICAL STANDARDS

Funding

No sources of funding were used to assist in the preparation of this article.

Conflicts of interest

This paper represents part of the work by the World Anti-Doping Agency Working Group on Doping Prevalence conducted between September 2017 and December 2019, but WADA had no control over the drafting or content of this manuscript. John Gleaves, Andrea Petróczi, Olivier De Hon, Martial Saugy and Maarten Cruyff served as members of the Working Group (2017-2019) and they prepared this paper in their capacity as Working Group members, in collaboration with DF and EM. The Working Group members receive no salary for their work but expenses related to the travel for work were covered. Andrea Petróczi received grant funding from WADA previously as part of the Social Science Research Program, served as a member of the first Working Group on Doping Prevalence (2011-2012); and is currently involved in providing analysis and evaluation support for WADA's Outreach Program in an unpaid advisory role. Martial Saugy worked at the Swiss Laboratory for Doping analyses (LAD, Centre Hospitalier Universitaire Vaudois (CHUV), Lausanne, Switzerland) until 2016 and received funding from WADA Science Department prior to his involvement in this project. Olivier De Hon works for the National Anti-Doping Authority Netherlands. Dirk Folkerts and Emmanuel Macedo declare they have no conflicts of interest relevant to the content of this review.

Consent to participate

Not applicable

Availability of data and materials

The definitions, questions, and rater criteria for the Quality Assessment of Doping evidence – Self-Reported Prevalence (QUAD-SRP) and the Quality Assessment of Doping evidence – Sample Analysis of Prevalence (QUAD-SRP) are available in Supplementary Material Appendix S1 and S2, respectively. All extracted data from the studies is available in Supplementary Material Appendix S3. The complete scoring for all studies is available in Supplementary Material Appendix S4. All other datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authorship contributions

Andrea Petróczi served as senior author on the project, conceptualized the study, led the development of quality assessment criteria, contributed to collating and synthesizing the independent quality assessments, contributed to the literature search, supervised Dirk Folkerts and contributed to drafting the manuscript. John Gleaves drafted the manuscript, contributed to the development of quality assessment criteria, contributed to collating and synthesizing the independent quality assessments as well as the literature search and supervised Emmanuel Macedo. Folkerts conducted the initial literature search, contributed to developing the quality assessment criteria and conducted independent quality assessment for all included studies under the supervision of Petróczi. Olivier De Hon conducted independent quality assessment, contributed to developing the quality assessment criteria and literature search. Macedo conducted independent quality assessment under the supervision of Gleaves. The best practice recommendations were formulated by Petróczi, Gleaves, De Hon, Martial Saugy, and Maarten

Cruyff. All authors read and critically commented on the manuscript and approved the final version of the manuscript.

Supplementary materials:

Supplementary Material Appendix S1: Quality Assessment of Doping evidence – Self-Reported Prevalence (QUAD-SRP)

M = Method and study design, E = Execution, D = Data analysis; R = Reporting

Data Quality						
Code	**Q#**	**Component**	**Definition**	**Question**	**Criteria**	**Score**
M	1	External validity	Addresses the extent to which the findings from the study can be generalised to the population from which the study subjects are derived.	Was the reported sampling method representative of the population intended to the study?	[A] Probability sampling (including: simple random, systematic, stratified g, cluster, two-stage and multi-stage sampling)	1
					[B] Entire population is invited to participate	1
					[C] Non-probability sampling (including: purposive, quota, convenience and snowball sampling) or no method reported	0
M	2	External validity	Addresses the extent to which survey used validated instruments standardized to the field.	Was the measure used to establish evidence for doping behaviour validated or aligned to field specific standards (e.g., epidemiology studies on substance use; or the WADA Research Pack)?	[A] Yes	1
					[B] Partially (use format from previous studies)	0.5
					[C] No (use bespoke format or not reported)	0
E	3	External validity	Addresses the extent to which the sample size is appropriate for establishing doping prevalence.	Is the sample size adequate for establishing prevalence? If direct questions used:	[A] Over 250	1
					[B] 50 – 250	0.5
					[C] Less than 50	0
				If indirect estimations (e.g., randomised response or fuzzy response models) used:	[A] Over 500	1
					[B] 250 – 500	0.5
					[C] Less than 250	0
M	4			Is their reported	[A] Yes	1

		Internal validity	Considers data quality in terms of whether method is appropriate to obtain evidence for doping behaviour.	evidence that the method executed was suitable for obtaining evidence for doping behaviour?	[B] No	0
M	5	Internal validity	Assesses whether the measurement of doping was objective.	How was the doping behaviour established?	[A] Self-reported	1
					[B] Inferred at individual level (, reaction time measures, performance, etc.)	0.5
					[C] Inferred at population level (e.g., trends in performance records, wastewater analysis, prescription rates, recorded purchase, police seizure)	0
M	6	Internal validity	Assess whether participants understood what was meant by doping when answering questions.	Did study report providing participants with a clear definition of doping at the time of data collection?	[A] Yes	1
					[B] No (including if unreported)	0
M	7	Internal validity	Assess whether participants understood the time frame for doping behaviour being measured.	Did study report providing participants a clear time frame (e.g. last 12 months) for their self-reported doping behaviour at the time of data collection?	[A] Yes	1
					[B] No (including if unreported)	0
M	8	Internal validity	Considers clarity of instructions and assurance for non-exposure (for indirect estimation methods only).	Did the study have mechanism in place to ensure that participants are not exposed?	[A] Yes	1
					[B] No (including if unreported)	0
					[C] Not applicable	NA
BIAS						

M	9	Framing	Addresses bias in the measurement of the outcomes in a study.	Was the reported doping question neutral and factual about the of behaviour being self-reported? (i.e., making no value-judgement about doping use in the question) Can the inferred 'use' be linked directly to the evidence (e.g., performance profiling)?	[A] Yes	1
					[B] No (including if unreported)	0
					[C] Not applicable	NA
M	10	Substance	Addresses bias in the interpretation of ambiguous substances (e.g., cannabis, alcohol, prescription drugs).	Did the reported doping question clearly differentiate prohibited doping behaviour from recreational or therapeutic use of prohibited substances?	[A] Yes	1
					[B] No (including if unreported)	0
					[C] Not applicable	NA
E	11	Sample	Addresses potential bias in data collection procedure through loss of control.	Did study report controlling data collection to ensure the integrity or the sample/answer (i.e., ensured that the participant is the athlete whose doping behaviour is assessed)?	[A] Yes	1
					[B] No (include postal surveys and internet surveys unless a unique ID code is used)	0
					[C] Not applicable	NA
D	12	Confounding factors	Addresses whether studies have applied adjustment for confounding in the analysis.	Did the study report controlling for confounding factors (e.g. stratification/ matching/ restriction/ adjustment) when analyzing the associations (if the study contains purely descriptive results based on direct survey so no association and prediction tests were conducted in	[A] Yes	1
					[B] No	0
					[C] Not applicable	NA

				the test, please select "Not applicable")?		
DA	13	Exclusion	Considers whether the dataset was altered retrospectively (i.e., altered after data collection completed, during data analysis).	If some participants were excluded post data collection stage, did study explain reason for excluding the participants and offer relevant demographics on the final dataset?	[A] Yes	1
					[B] No	0
					[C] Not applicable	NA
EX	14	Privacy and confidentiality	Considers whether participants were sufficiently protected during data collection.	Did study report protecting privacy or sensitivity of the nature of the doping behaviour considered when data were collected?	[A]Yes	1
					[B] No	0
R	15	Reporting	Assesses whether the information provided in the paper is sufficient to allow a reader to make an unbiased assessment of the findings of the study.	Did the study report any response rate?	[A] Above 70%	1
					[B] 50 to 70%	0.5
					[C] Not reported or below 50% is 0	0
RP	16	Reporting	Assess whether the study controlled for non-compliant responses.	Did the study report the rate non-compliance (i.e., participants refused to answer the doping question or were otherwise non-compliant)?	[A] Yes	1
					[B] No	0
RP	17	Reporting	Assess whether study reports accurate evidence for doping behaviour.	Did the study report clearly how evidence for doping behaviour was established? If only partial information is	[A] Yes	1
					[B] No	0

				provided, answer 'no'.		

[A] Total Points Available	17
[B] Total Points Earned (Sum of Points 1-17)	
[C] Total N/A	
[D] Score % [B/(A-C)]	

Supplementary Material Appendix S2: Quality Assessment of Doping evidence – Sample Analysis of Prevalence (QUAD-SAP)

M = Method and study design, E = Execution, D = Data analysis; R = Reporting

Data Quality						
Code	**Q#**	**Component**	**Definition**	**Question**	**Criteria**	**Score**
M	1	External validity	Addresses the extent to which the findings from the study can be generalised to the population from which the study subjects are derived.	Was the reported sampling method representative of the population intended to the study? (what constitute 'population' is defined by the researchers)	[A] Sample analysis included all available samples	1
					[B] Sample analysis included entire population	1
					[C] Sample analysis was performed for only a portion of the identified population	0
M	2	External validity	Addresses the extent to which sample analysis can indicate population.	Did the sample analysis control for the number of athletes tested?	[A] Yes	1
					[B] Partially (indicates number of athletes but does not consider an athlete providing multiple positive tests)	0.5
					[C] No (simply looks at total samples provided)	0
E	3	External validity	Addresses the extent to which the sample size is appropriate for establishing doping prevalence.	Is the data set size adequate for establishing prevalence? If examining athletes that provided samples?	[A] Over 250 athletes	1
					[B] 50 – 250 athletes	0.5
					[C] Less than 50 athletes	0
				If examining samples that do not account for athletes?	[A] Over 750 samples	.5
					[C] Less than 750 samples	0
E	4	Internal validity	Addresses the extent to which the reported findings are appropriate for establishing doping prevalence.	Does the reported screening of the sample analysis control for a sample having multiple prohibited substances?	[A] Yes, reports number of athletes screened that provided a positive sample for analysis	1
					[B] No, reports total number of prohibited substances identified, but does not control for an athlete providing a sample positive for multiple prohibited substances.	0
M	5	Internal validity	Considers data quality in terms of whether method is	Did the sample analysis screen for the full menu of prohibited substances? (If using	[A] Yes	1
					[B] Partially (only screened for a limited number of prohibited substances)	.5

			appropriate to obtain evidence for doping behaviour.	WADA approved laboratory screening, mark yes)	[B] No	0
M	6	Internal validity	Assess whether participants were competitive athletes.	Did study only screen samples from athletes in competitive sport?	[A] Yes	1
					[B] No (including if unreported)	0
M	7	Internal validity	Assess whether the time frame for doping behaviour was measured.	Did study report the time frame (e.g. last 12 months) for the analysis of samples.	[A] Yes	
					[B] No (including if unreported)	0
DA	8	Exclusion	Considers whether the dataset included samples that were not doping behaviors.	Did sample analysis control for therapeutic use exemptions, atypical finds, or other possibilities that were not anti-doping rule violations.	[A] Yes (only reported analyses qualifying as ADRVs)	1
					[B] Partial (removed some analyses not qualifying as ADRVs)	.5
					[C] No (did not control for TUE or ATF)	0
DA	9	Exclusion	Considers the extent of inadvertent doping.	Did sample analysis separate deliberate from accidental doping?	[A] Yes	1
					[B] No	0
M	10	Internal validity	Considers the confounding factors when prevalence of doping is inferred for population level.	Did the estimation take confounding factors into account? (e.g., drugs present in the water before 'athletes' samples'; travel, living or lifestyle conditions for ABP samples)	[A] Yes	1
					[B] No	0
					[A] Total Points Available	10
					[B] Total Points Earned (Sum of Points 1-10)	
					[C] Total N/A	
					[D] Score % [B/(A-C)]	

Supplementary Material Appendix S3: Summary Table of Evidence Survey

REF	Reference	Pub Year	Study GP	Study Method	Prevalence	Date for Prevalence	Participants in study			Samples in study			Country	Survey Question to Establish Use	Timeframe for doping behavior
							Study Parts. (N)	Men (N)	Women (N)	Sample (N)	Men Sample (N)	Women Sample (N)			
23	Afolayan (2012)	2012	SRP	DS	3.60%	NR	220	135	85				Nigeria	No exact wording information available.	Current and lifetime use
112	Aguilar et al. (2017)	2017	SAP	TF	1.9%	2003-2015	NR	NR	NR	NR	NR	NR	International	NA	NA
113	Aguilar-Navarro et al. (2019)	2019	SAP	TF	0.9-3.3%	2003-2015	NR	NR	NR	1347213	NR	NR	International	NA	NA
24	Ajayi-Vincent & Olanipekun (2014)	2014	SRP	DS	12.94-22.75%	NR	510	255	255				Nigeria	No exact wording information available.	Not Indicated
25	Al Ghobain (2017)	2017	SAP	TF	1-6.6%	2008-2016	NR	NR	NR	4482	NR	NR	Saudi Arabia	NA	NA
114	Al Ghobain et al. (2016)	2016	SRP	DS	4.3%	NR	1142	1142	0				Saudi Arabia	"Ever used any type of prohibited substances''	Lifetime
26	Ama et al. (2003)	2003	SRP	DS	7.0%	NR	1116	1037	79				Cameroon	"Do you use cocaine before matches?"	Not Indicated
27	Anderson et al. (1991)	1991	SRP	DS	1-10% (AAS only)	1985, 1989	3264	2150	1114				United States	No exact wording information available.	Current user and when athlete started
125	Aubel et al. (2019)	2019	SAP	AD	2.6%	2005-2016	NR	NR	0	NR	NR	NR	International	NA	NA
28	Backhouse et al. (2013)	2013	SRP	DS	13%	NR	212	137	75				United Kingdom	No exact wording information available.	Not Indicated
115	Bahr & Tjørnhom (1998)	1998	SAP	TF	1.2% Norway; 2.1% International	1977-1995	NR	NR	NR	15208	11931	3277	Norway	NA	NA
29	Barkoukis et al. (2011)	2011	SRP	DS	8.0%	NR	1040	651	389				Greece	"Have you ever used PES [Performance Enhancing Substances]"	Once, occasional, or systematic
32	Barkoukis et al. (2013)	2013	SRP	DS	9.9%	NR	750	479	271				Greece	"Have you ever used prohibited substances to enhance performance?''	Lifetime
30	Barkoukis et al. (2015)[a]	2015	SRP	DS	4.2%	NR	643	NR	NR				Greece	No exact wording information available.	Once
31	Barkoukis et al. (2020)	2019	SRP	DS	RS	NR	497	318	179				NR	"Have you ever used prohibited substances or methods to enhance your performance?)"	Lifetime
33	Boardley et al (2017)	2017	SRP	DS	8.20%	NR	364	223	141				United Kingdom	Participants were provided with a list of nine categories of doping substances (e.g., Ephedrine stimulants) and methods (e.g., Blood manipulation) based on WADA's prohibited list and to indicate time frame for use.	Currently used, used in the last 3 months, used prior to last three months, or never used.
102	Boardley et al (2019)	2019	SRP	RT	13.90%	NR	822	532	290				International	"Have you knowingly used substances [eg, anabolic steroids, erythropoietin, banned stimulants, growth hormones] or methods [eg, blood infusions] during the past 12 months that are banned by the WADA and the IOC and therefore would not be permitted in professional sport?"	Past 12 months
34	Buckman et al. (2009)	2009	SRP	DS	31.0%	NR	234	234	0				United States	''Within the last 12 months, have you taken [substances listed]''	Past 12 months

35	Buckman et al. (2013)	2013	SRP	DS	3.1%	2008-2009	11556	11556	0				United States	"Please indicate your experience with the following substances: [substances listed]''	Never Used, Used in the last 30, days, Used in the last 12 months, or Used, but not in the last 12 months.
36	Chiang et al. (2018)	2018	SRP	DS	4.40%	2017	182	109	73				Malaysia	No exact wording information available.	Not Indicated
37	Coopoo et al. (2000)	2000	SRP	DS	NR	1998	140	74	66				South Africa	No exact wording information available.	Not Indicated
38	Corbin et al. (1994)	1994	SRP	DS	1.9%	NR	1690	1013	677				United States	"I have tried anabolic steroids.?" (Yes/No)	Lifetime
39	Curry & Wagman (1999)	1999	SRP	DS	46.66% Past Year; 66.7% Lifetime	1986-1991	15	15	0				United States	No exact wording information available.	Lifetime
40	Da Silva (2017)	2017	SRP	DS	.05-1.7%	2006	402	197	205				Brazil	No exact wording information available.	Not Indicated
41	Dezelsky et al. (1985)	1985	SRP	DS	15%-20%	1970-1984	4171	NR	NR				United States	No exact wording information available.	Not Indicated
104	Dietz et al. (2013)	2013	SRP	RT	13.0%	NR	2987	2576	376				Germany	Have you used substances which can only be prescribed by a doctor, are available in a pharmacy, or can be bought on the black market (e.g. anabolic steroids, erythropoietin, stimulants, growth hormones) to enhance your physical performance during the last 12 months?"	Past 12 months
103	Dietz et al. (2014)	2014	SRP	DS	0.8%	NR	525	317	208				Germany	"Have you used substances which can only be prescribed by a doctor, are available in a pharmacy, or can be bought on the black market (e.g. caffeine tablets, stimulants, cocaine, methylphenidate, modafinil, beta blockers) to enhance your performance during the last 12 months?''	Past 12 months
42	Dietz et al. (2016)	2016	SRP	RT	12.4-20.4%	NR	2702	2358	344				Germany	No exact wording information available.	Not Indicated
43	Dodge et al. (2008)	2008	SRP	DS	2.50%	NR	241	154	81				United States	"Have you ever tried an illegal performance enhancing substance?''	Lifetime
44	Dodge et al. (2013)	2013	SRP	DS	0.0%	NR	132	132	0				United States	"Have you ever used a substance to help improve your athletic/physical performance?''	Lifetime
45	Donahue et al. (2006)	2006	SRP	DS	RS	NR	1201	637	650				Canada	"Have you used…[Listed substances]'' with 5-point Likert scale, ranging from 0 (No) to 4 (Yes, I use it regularly).	Not Indicated

105	Elbe & Pitsch (2018)	2018	SRP	RT	0-30.6% Past Season; 3.1-26.0% Lifetime	NR	771	435	336				Denmark	"Have you ever intentionally used forbidden substances and/or for-bidden methods in order to enhance your sporting performance in competitions?" and "Have you intentionally used forbidden substances and/or forbidden methods in order to enhance your sporting performance in competitions during the last season?"	Lifetime and during the last season
46	Elliot et al. (2004)	2004	SRP	DS	0.0-0.3%	NR	668	0	668				United States	No exact wording information available.	Lifetime use and use during the season
101	Erickson et al (2019)	2019	SRP	DS	1-3%	NR	568	301	267				International	Participants indicated the frequency they had taken each substance by circling one of the options: "never", "once a month", "once a week", "more than once a week" or "don't know".	Last 3 months
122	Faiss et al. (2019)	2019	SAP	BP	18%	2011, 2013	1222	700	522	3683	2008	1675	International	NA	NA
47	Gallucci et al. (2015)	2015	SRP	DS	7.5%	2014	200	78	122				United States	No exact wording information available.	Last 12 months
48	Goulet et al. (2010)	2010	SRP	DS	25.8%	NR	3573	2000	1573				Canada	No exact wording information available.	Last 12 months
49	Gradidge et al. (2011)	2011	SRP	DS	4-5%	NR	100	100	0				South Africa	No exact wording information available.	Not Indicated
50	Green et al. (2001)	2001	SRP	DS	1.10%	NR	13914	9183	4722				United States	No exact wording information available.	Last 12 months
51	Gucciardi et al. (2011)	2011	SRP	DS	0.8%	NR	643	285	383				Australia	"Which one of the following most applied to you? Items: never, briefly, moderately, still think about it, briefly used, occasionally used, regularly use"	Not Indicated
52	Hejabi et al. (2015)	2015	SRP	DS	RS	NR	373	176	197				India	For list of 20 prohibited substances, asked Likert 5-value scale from zero for "No, I don't use" up to 5 with "yes, I usually use."	Last 12 months
53	Horn et al. (2009)	2009	SRP	DS	9.1%	2001-2003	2552	2552	0				United States	'During the time in which it was acceptable to use performance-enhancing steroids, did you use steroids?'	Specific window of years
54	Jalleh et al. (2014)	2014	SRP	DS	3.4-6.9%	NR	1237	602	635				Australia	"In the last 12 months, have you used any of the following, for whatever reason: (List) Response: (1) Have never used; (2) Did not use in the last 12 months; (3) 1 to 2 times; (4) 3 to 5 times; (5) 6 to 10 times; (6) More than 10 times.	Last 12 months

55	James et al (2013)	2013	SRP	DS,NS	19.88%-58.41%	NR	513	301	212				United Kingdom	"I have taken prohibited performance enhancing drugs in the past 12 months." (Yes/No)	Last 12 months
56	Judge et al. (2012)	2012	SRP	DS	3.4%	NR	98	46	52				United States	''Have you ever used prohibited substances to enhance your performance?''	Lifetime
57	Kabiri et al. (2018)[b]	2018	SRP	DS	RS	NR	606	365	241				Iran	Participants were asked to report whether they (a) "currently use a banned substance," (b) had "previously used a banned substance to enhance their performance," or (c) "intended to use a banned substance at least once within the next twelve months," with response rating of 0 (never) to 3 (systematically) used banned substances.	Current and lifetime use
58	Kabiri et al. (2019)[b]	2019	SRP	DS	1.10%	NR	852	494	358				Iran	"Currently use a banned substance"	Current use
59	Kersey (1996)	1996	SRP	DS	3.30%	NR	1185	833	352				United States	No exact wording information available.	Not Indicated
60	Kisaalita & Robinson (2014)	2014	SRP	DS	7.4%	NR	68	61	7				United States	"Cyclist responded either "yes" or "no" as to whether they used non-banned PEPs"	Not Indicated
61	Krowchuk et al. (1989)	1989	SRP	DS	1-2%	1987-1988	295	212	83				United States	No exact wording information available.	Not Indicated
123	Kuipers et al. (2007)	2007	SAP	BP	0-2%	2000-2005	NR	NR	NR	975	556	419	International	NA	NA
62	Lazuras et al. (2010)	2010	SRP	DS	9.90%	NR	750	479	271				Greece	Have you ever used PES [Performance Enhancing Substances]	Lifetime, once, occasional, systematic
63	Lazuras et al. (2015)[a]	2015	SRP	DS	4.2%	NR	650	444	206				Greece	How frequently do you use PES to improve your performance?	Not Indicated
64	Lindqvist et al. (2013)	2013	SRP	DS	21.0%	NR	683	683	0				Sweden	No exact wording information available.	Lifetime
65	Ljungqvist (1975)	1975	SRP	DS	31%	1973	99	99	0				Sweden	No exact wording information available.	Not Indicated
66	Lorente et al. (2005)	2005	SRP	DS	1.8-10.7%	NR	1152	665	487				France	No exact wording information available.	Lifetime
67	Manouchehri & Tojari (2013)[c]	2013	SRP	DS	11.9%	NR	160	120	40				Iran	''Do you currently use banned performance-enhancing drugs?'' / ''Have you ever had personal experience with banned performance-enhancing drugs and/or methods?''	Current and lifetime use
68	Manouchehri & Tojari (2013)[c]	2013	SRP	DS	11.9%	NR	160	120	40				Iran	''Do you currently use banned performance-enhancing drugs?'' / ''Have you ever had personal experience with banned performance-enhancing drugs and/or methods?''	Current and lifetime use

126	Maquirriain (2010)	2010	SAP	AD	0.38%	2003-2009	NR	NR	NR	13340	NR	NR	International	NA	NA
116	Marchand et al (2017)	2017	SAP	TF	16.67%	NR	42	42	0	42	42	0	Guadeloupe	NA	NA
117	Mazzeo et al (2016)	2016	SAP	TF	2.7-4.75%	2003-2010	NR	NR	NR	15396	10347	5049	Italy	NA	NA
118	Mercado et al. (2019)	2019	SAP	TF	3.8%	2009-2015	NR	NR	NR	18085	NR	NR	Mexico	NA	NA
69	Molobe (2012)	2012	SRP	DS	.09-6.2%	NR	345	208	137				Nigeria	No exact wording information available.	Not Indicated
111	Morente-Sanchez et al. (2013)	2013	SRP	QI	0.0%	2012	8	0	8				Spain	''Have you ever used doping substances?''	Lifetime
70	Morente-Sánchez et al. (2019)	2019	SRP	DS	4.5%	2012-2013	1324	1276	48				Spain	"The use of supplements and which ones are taken?"	Not Indicated
71	Mudrak et al. (2018)	2018	SRP	DS	1.18%	2014-2015	1035	667	368				Czech Republic	No exact wording information available.	Not Indicated
72	Muwonge et al. (2015)	2015	SRP	DS	3.3% Current; 3.9% Lifetime	NR	360	218	142				Uganda	No exact wording information available.	Not Indicated
73	Naylor et al. (2001)	2001	SRP	DS	2.5%	NR	1121	NR	NR				United States	"Have you violated this rule during the season?"	Past Season (undefined)
74	Nica-Badea (2014)[d]	2014	SRP	DS	1.1%	NR	171	NR	NR				Romania	No exact wording information available.	Not Indicated
75	Ntoumanis et al (2017)	2017	SRP	DS	16.90%	NR	166	NR	NR				Greece	"Presented athletes with a list of five of the most common doping substances (i.e., testosterone and byproducts, growth hormone and IGF-1, β-blockers, erythropoietin, and anabolic steroids)." Participants responded in a yes–no format.	Past 6 months
76	Ogama et al. (2019)	2019	SRP	DS	NR	NR	291	NR	NR				Kenya	No exact wording information available.	Not Indicated
77	Ohaeri et al (1993)	1993	SRP	DS	1.20%	1992	250	180	70				Nigeria	"Have you ever used the drugs?" (of a list attached)	Lifetime use and frequency
78	Özdemir et al. (2005)	2005	SRP	DS	14.50%	NR	433	350	83				Turkey	No exact wording information available.	Not Indicated
119	Pereira & Sardela (2014)	2014	SAP	TF	1.6-2.1%	2003-2011	NR	NR	NR	43478	NR	NR	Brazil	NA	NA
79	Petroczi (2007)	2007	SRP	DS	2.5% Current; 7.5% Past Use	NR	199	199	0				United States	''Have you ever had personal experience with banned performance-enhancing drugs and/or methods?'' / ''Do you currently use banned performance-enhancing drugs?''	Lifetime and current use
80	Petroczi et al. (2011)	2011	SRP, SAP	DS, HS	12.2-13.4%	NR	82	39	43	82	39	43	Hungary	''Have you ever used a banned substance?''	Lifetime
108	Pitsch & Emrich (2011)	2011	SRP	RT	9.6-20.4 Current Season; 10.2-25.8% Lifetime Use	2005-2008	5409	NR	NR				Germany	''Have you ever…?'' / ''Have you…in the current season?''	Lifetime and current season
106	Pitsch (2018)	2018	SRP	RT	3.15%	2014	2949	NR	NR				United States	Have you used forbidden substances or methods in order to enhance your cycling performance during the last season? (TUE excluded)	Last season (undefined)
107	Pitsch at al. (2007)	2007	SRP	RT	20.4-48.1%	NR	448	296	152				Germany	''Have you ever…?'' / ''Have you…in the current season?''	Lifetime and current season
81	Rodek et al. (2009)	2009	SRP	DS	67.0%	NR	27	27	0				Bosnia and Herzegovina	Select: ''I do not use doping, / I use doping from time to time / I	Not Indicated

														use doping on a regular basis.''	
82	Ruzdija et al. (2018)	2018	SRP	DS	Under 18 years of age: 7.7% Current; 19.2% Lifetime Over 18 years of age: 22.7% Current; 22.7% Lifetime	NR	48	NR	NR				Macedonia	''Have you ever had personal experience with banned performance-enhancing drugs and/or methods?'' / ''Do you currently use banned performance enhancing drugs?''	Lifetime and current use
83	Sánchez-Oliver et al. (2019)	2019	SRP	DS	22.9-72.%	NR	48	44	4				Spain	No exact wording information available.	Not Indicated
84	Schneider et al. (1993)	1993	SRP	DS	4.6%	NR	197	142	55				United States	No exact wording information available.	Not Indicated
109	Seifarth et al. (2019)	2019	SRP	RT	7.0%	NR	1989	1477	456				Germany	"Have you taken substances to increase your physical performance within the past 12 months that are only available at a pharmacy, at the doctor's office, or on the black market (e.g., anabolic steroids, erythropoietin, stimulants, growth hormones)?"	Last 12 months
85	Silvester (2006)	2006	SRP	DS	61-68%	1972	100	100	0				International	"Have you taken anabolic steroids within the past six months? Have you ever taken steroids?"	Last 6 months and lifetime
86	Soltanabadi et al. (2015)[e]	2015	SRP	DS	NR	NR	200	114	86				Iran	No exact wording information available.	Not Indicated
124	Sottas et al. (2011)	2011	SAP	BP	1-48%	2001-2009	2737	NR	NR	7289	4009	3280	International	NA	NA
87	Stilger & Yesalis (1999)	1999	SRP	DS	6.30%	NR	873	873	0				United States	No exact wording information available.	Lifetime and current use
88	Striegel et al. (2010)	2010	SRP	DS, RT	0.2-6.8%	NR	1458	912	543				Germany	"Have you ever used doping substances"	Lifetime
89	Tahtamouni et al. (2008)	2008	SRP	DS	26%	NR	154	NR	NR				Jordan	No exact wording information available.	Past and current use
90	Terney & McLain (1990)	1990	SRP	DS	5.50%	1988	1436	833	603				United States	Have you ever used anabolic steroids?	Lifetime
91	Tricker & Connolly (1997)	1997	SRP	DS	0.32% Current; 3.6% Lifetime	NR	635	435	200				United States	No exact wording information available.	Not Indicated
92	Uduwana & Madushani (2014)	2014	SRP	DS	65%	NR	60	47	13				Sri Lanka	No exact wording information available.	Not Indicated
110	Ulrich et al. (2018)	2018	SRP	RT	43.6%- 57.1%	2011	2167	NR	NR				International	"Have you knowingly violated anti-doping regulation by using a prohibited substance or method in the past 12 months?"	Last 12 months
93	Uvacsek et al. (2011)[f]	2011	SRP	DS	14.63%	NR	82	37	45				Hungary	No exact wording information available.	Not Indicated
94	Vajiala et al. (2010)	2010	SRP	DS	0.01%	NR	1404	NR	NR				Romania	No exact wording information available.	Not Indicated
120	Van Eenoo & Delbeke (2003)	2003	SAP	TF	5.8%	1993-2000	NR	NR	NR	14995	13224	1771	Belgium	NA	NA
121	Vouillamoz et al. (2009)	2009	SAP	TF	0.0%	2008	286	286	0	286	286	0	International	NA	NA
95	Wagman et al. (1995)	1995	SRP	DS	66.70%	1986-1991	15	15	0				United States	No exact wording information available.	Lifetime
96	Wanjek et al. (2007)	2007	SRP	DS	0.2%-0.9%	NR	1751	886	865				Germany	List of substances with 7 point rating scale between "at no	Last 12 months

														time" and "at all time"	
127	Whitaker & Backhouse (2017)	2017	SAP	AD	NR	2009-2014	NR	NR	NR	NR	NR	NR	United Kingdom	NA	NA
97	Whitaker et al. (2014)	2014	SRP	DS	2.3% Current; 4.5% Lifetime	NR	729	460	269				United Kingdom	Asked whether "they currently used a banned substance, whether they had previously used a banned substance to enhance their performance and whether they intended to use a banned substance at least once within the next 12 months."	Current and last 12 months
98	Wroble et al. (2002)	2002	SRP	DS	0.70%	NR	1553	1087	466				United States	''Have you ever used anabolic steroids?''	Lifetime
99	Yesalis (1988)	1988	SRP	DS	24%-55%	NR	45	NR	NR				United States	"Have you ever used anabolic steroids?"	Lifetime
100	Zenic et al. (2010)	2010	SRP	DS	5%	NR	38	0	38				Croatia	No exact wording information available.	Not Indicated

Self-reported prevalence (SRP); Sample analysis for prevalence (SAP); Direct survey (DS); Random response technique (RT); Testing figure (TF); Athletes biological passport (BP); Anti-doping rule violation (AD); Qualitative interview (QI); Hair sample analysis (HS); Network scale-up (NS); Not Reported (NR); Reported as a Scale (RS); Anabolic Androgenic Steroid (AAS). Studies marked with superscript letters contain identical datasets.

Notes: [d]: also published in Nica-Badea, D. (2016). Social Determinants of Intention to Dope in Sports Clubs and Institutions. Annals of Applied Sport Science, 4(2), 33-40; [e]: also published in Tojari, F., Manouchehri, J., & Soltanabadi, S. (2015). Examining Conceptual Model of the Relationships between Sports Motivation, Doping Attitudes and Doping Behavior in Professional Athletes. Journal of Applied Environmental Biological Sciences, 5(7), 305-310. . These papers were not included owing to unclear reporting on prevalence.[f]: a subset of the data were previously published in Petróczi, A., Aidman, E. V., Hussain, I., Deshmukh, N., Nepusz, T., Uvacsek, M., Tóth, M., Barker, J., & Naughton, D. P. (2010). Virtue or pretense? Looking behind self-declared innocence in doping. PloS one, 5(5), e10457.

Supplementary Material Appendix S4: Scores for both study groups

QUAD-SRP Scores for Self-Reported Prevalence

Study GP	#	Articles	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	Rating
SR	23	Afolayan (2012)	1	1	0.5	1	1	0	0	NA	0	0	1	0	NA	1	0	0	1	50%
SR	24	Ajayi-Vincent & Olanipekun (2014)	1	1	1	1	1	1	0	0	1	1	0	0	NA	0	1	0	1	63%
SR	114	Al Ghobain et al. (2016)	1	1	1	1	1	0	0	NA	1	0	1	0	NA	1	1	0	1	67%
SR	26	Ama et al. (2003)	0	0	1	0	1	0	0	NA	1	0	0	1	NA	1	1	0	1	47%
SR	27	Anderson et al. (1991)	1	0.5	1	1	1	0	1	NA	1	0	1	NA	NA	1	0	0	1	68%
SR	28	Backhouse et al. (2013)	0	1	0.5	1	1	1	0	NA	1	0	0	1	NA	0	0	0	1	50%
SR	29	Barkoukis et al. (2011)	1	1	1	1	1	0	1	NA	1	0	1	0	1	1	0.5	1	1	78%
SR	32	Barkoukis et al. (2013)	1	1	1	1	1	0	1	NA	1	0	1	0	NA	1	0	0	1	67%
SR	33	Barkoukis et al. (2015)	1	1	1	1	1	1	1	NA	1	1	1	0	NA	1	1	0	1	87%
SR	31	Barkoukis et al. (2020)	0	0	1	1	1	1	0	NA	1	1	1	0	NA	1	1	0	1	67%
SR	33	Boardley et al (2017)	0	0.5	1	1	1	1	1	NA	1	0	0	0	NA	1	0	0	1	57%
SR	102	Boardley et al (2019)	0	1	1	0.5	1	1	1	0	1	1	0	0	NA	1	0	0	1	59%
SR	34	Buckman et al. (2009)	0	0.5	0.5	1	1	1	1	NA	1	1	1	1	1	1	0	1	1	81%
SR	35	Buckman et al. (2013)	1	1	1	1	1	1	1	NA	1	1	1	1	1	1	0	0	1	88%
SR	36	Chiang et al. (2018)	0	1	0.5	1	1	0	0	NA	0	1	0	NA	1	0	0	1	1	50%
SR	37	Coopoo & Jakoet (2000)	1	1	0.5	1	1	0	1	NA	1	0	0	NA	NA	1	0.5	0	1	64%
SR	38	Corbin et al. (1994)	0	0	1	1	1	1	1	NA	1	1	1	0	1	1	1	0	1	75%
SR	39	Curry & Wagman (1999)	0	0	0	1	1	1	1	NA	1	1	0	NA	NA	1	0.5	1	1	68%
SR	40	Da Silva (2017)	1	0	1	1	1	0	0	NA	0	1	0	0	NA	0	0	0	1	40%
SR	41	Dezelsky et al. (1985)	1	0.5	1	1	1	0	1	NA	0	0	0	NA	NA	0	0	0	0	39%
SR	104	Dietz et al. (2016)	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	88%
SR	103	Dietz et al. (2013)	1	1	1	1	1	1	1	0	1	1	1	1	NA	1	1	0	1	88%
SR	42	Dietz et al. (2014)	1	0.5	0.5	1	1	1	1	NA	1	0	1	1	0	1	1	0	1	75%
SR	43	Dodge et al. (2008)	1	0	0.5	1	1	1	1	NA	1	1	1	1	NA	1	0	0	1	77%
SR	44	Dodge et al. (2013)	1	0	0.5	1	1	0	1	NA	1	0	1	1	NA	1	0	0	1	63%
SR	45	Donahue et al. (2006)	1	1	1	1	1	0	1	NA	1	0	0	0	1	1	0	0	1	63%
SR	105	Elbe & Pitsch (2018)	1	1	1	1	1	0	1	1	0	0	1	0	NA	1	0.5	1	1	72%
SR	46	Elliot et al. (2004)	1	0	1	0	1	0	0	NA	0	0	1	0	1	1	1	0	0	44%
SR	101	Erickson et al (2019)	0	0	1	1	1	1	1	NA	1	0	1	NA	1	1	0	0	1	67%

SR	47	Gallucci & Martin (2015)	0	1	1	0	1	0	1	NA	0	0	1	0	1	1	1	0	1	56%
SR	48	Goulet et al. (2010)	0	1	1	1	1	0	1	NA	1	0	0	0	NA	0	0	0	0	40%
SR	49	Gradidge et al. (2011)	1	0.5	0.5	1	1	0	1	NA	1	0	1	NA	0	1	1	0	1	67%
SR	50	Green et al. (2001)	1	0.5	1	1	1	0	1	NA	1	0	1	1	NA	1	0.5	0	0	67%
SR	51	Gucciardi et al. (2011)	0	0	0.5	0.5	1	0	1	NA	1	0	0	1	1	1	0	0	0	44%
SR	52	Hejabi et al (2015)	1	0.5	1	1	1	1	1	NA	1	1	1	NA	NA	0	0	0	0	68%
SR	53	Horn et al. (2009)	1	0.5	1	1	1	1	1	NA	0	0	0	NA	NA	1	0.5	0	1	64%
SR	54	Jalleh et al. (2014)	1	0.5	1	1	1	0	1	NA	0	0	0	NA	1	1	0	1	1	63%
SR	55	James et al. (2013)	1	0.5	0.5	1	1	0	1	NA	1	0	1	NA	NA	1	0	0	1	64%
SR	56	Judge et al. (2012)*	0	0.5	0.5	1	1	0	1	NA	1	0	1	1	NA	1	0	0	1	60%
SR	57	Kabiri et al. (2018)	1	1	1	1	1	0	1	NA	0	0	0	0	NA	1	0	0	0	47%
SR	58	Kabiri et al. (2019)	1	0	1	1	1	0	0	NA	0	0	0	NA	NA	0	1	0	1	43%
SR	59	Kersey (1996)	1	1	1	1	1	1	0	NA	1	0	1	NA	NA	1	0	0	1	71%
SR	60	Kisaalita & Robinson (2014)	0	0	0.5	1	1	0	0	NA	0	0	0	NA	NA	1	0	0	1	32%
SR	61	Krowchuk et al. (1989)	1	0	1	1	1	0	0	NA	1	0	1	NA	NA	1	1	1	1	71%
SR	62	Lazuras et al. (2010)	1	0	0.5	1	1	1	1	NA	1	0	1	NA	1	1	0.5	0	1	73%
SR	63	Lazuras et al. (2015)	1	0.5	0.5	1	1	0	0	NA	1	0	1	1	NA	1	1	0	1	67%
SR	64	Lindqvist et al. (2013)	1	1	1	1	1	1	1	NA	1	0	0	NA	NA	0	0.5	0	1	68%
SR	65	Ljungqvist (1975)	0	0	0.5	1	1	0	0	NA	1	NA	NA	NA	NA	0	0.5	0	1	42%
SR	66	Lorente et al. (2005)	0	1	1	1	1	1	0	NA	1	1	1	NA	1	1	1	1	1	87%
SR	67	Manouchehri & Tojari (2013)a	0	1	0.5	1	1	0	0	NA	0	1	0	NA	NA	0	0	0	1	39%
SR	68	Manouchehri & Tojari (2013)b	1	0.5	0.5	0	1	0	0	NA	1	0	1	NA	NA	NA	0	0	0	38%
SR	69	Molobe (2012)	1	0	1	1	1	0	0	NA	1	0	1	NA	1	1	1	0	1	67%
SR	111	Morente-Sánchez et al. (2013)	0	0	0	1	1	0	0	NA	1	0	1	NA	NA	1	1	1	1	57%
SR	70	Morente-Sánchez et al. (2019)	0	1	1	1	1	0	1	NA	0	0	0	NA	0	1	0.5	0	1	50%
SR	71	Mudrak et al. (2018)	0	0	1	1	1	0	0	NA	0	1	0	NA	NA	1	1	0	1	50%
SR	72	Muwonge et al. (2015)	1	1	1	1	1	1	0	NA	1	0	1	NA	NA	1	1	0	1	79%
SR	73	Naylor et al. (2001)	1	0.5	1	1	1	1	1	NA	1	1	1	NA	NA	1	0	0	1	82%
SR	74	Nica-Badea (2014)	0	1	0.5	1	1	0	0	NA	0	NA	NA	NA	NA	NA	0	0	1	41%
SR	75	Ntoumanis et al (2017)	0	0.5	1	1	1	0	1	NA	1	0	1	1	NA	1	0	0	0	57%
SR	76	Ogama et al. (2019)	1	0.5	1	1	1	1	0	NA	1	1	1	NA	NA	1	0	0	0	68%
SR	77	Ohaeri et al. (1993)	0	0.5	0.5	1	1	0	1	NA	1	0	1	NA	NA	1	1	0	1	64%
SR	78	Özdemir et al. (2005)	1	1	1	1	1	1	0	NA	1	1	1	NA	NA	1	0	0	1	79%
SR	79	Petroczi (2007)	1	1	0.5	1	1	0	1	NA	1	0	1	1	1	1	0	0	1	72%

SR	80	Petroczi et al. (2011)	0	1	0.5	1	1	0	1	NA	1	0	1	NA	NA	1	0	0	1	61%
SR	108	Pitsch & Emrich (2011)	1	1	1	1	1	0	1	0	1	0	0	NA	NA	1	0	0	1	60%
SR	106	Pitsch (2018)	1	1	1	1	1	1	1	1	1	1	0	0	NA	1	0	1	1	81%
SR	107	Pitsch et al. (2007)	0	1	0.5	1	1	0	1	0	1	0	0	NA	0	1	0	0	1	47%
SR	81	Rodek et al. (2009)	0	0.5	0	1	1	0	1	NA	1	0	1	NA	NA	1	0	0	1	54%
SR	82	Ruzdija et al. (2018)	0	0.5	0	1	1	0	1	NA	1	0	0	NA	NA	0	0	0	1	39%
SR	83	Sánchez-Oliver et al. (2019)	0	0.5	0.5	1	1	0	0	NA	0	0	0	NA	NA	1	0	1	1	43%
SR	84	Schneider et al. (1993)	0	0	0.5	1	1	0	0	NA	0	0	0	NA	NA	1	0	0	1	32%
SR	109	Seifarth et al. (2019)	0	1	1	1	1	1	1	0	1	1	1	0	1	1	0	1	1	76%
SR	85	Silvester (2006)	0	1	1	1	1	1	1	NA	1	0	NA	NA	NA	0	0	0	1	62%
SR	86	Soltanabadi et al. (2015)	0	0.5	0.5	1	1	1	1	NA	1	0	1	NA	NA	0	0	0	1	57%
SR	87	Stilger & Yesalis (1999)	1	1	1	1	1	1	1	NA	1	0	1	NA	NA	1	0.5	0	1	82%
SR	88	Striegel et al. (2010)	1	1	0.5	1	1	0	1	0	1	NA	1	NA	NA	1	1	0	1	75%
SR	89	Tahtamouni et al. (2008)	0	1	0.5	1	1	1	1	NA	1	0	1	NA	NA	1	1	0	1	75%
SR	90	Terney & McLain (1990)	0	0	1	1	1	1	0	NA	1	0	1	NA	NA	1	0.5	0	1	61%
SR	91	Tricker & Connolly (1997)	0	0.5	1	1	1	0	1	NA	1	0	0	NA	NA	0	0	0	1	46%
SR	92	Uduwana & Madushani (2014)	0	1	0.5	0	1	0	0	NA	0	0	1	NA	NA	0	0	0	1	32%
SR	110	Ulrich et al. (2018)	1	1	1	1	1	1	1	1	1	1	1	NA	NA	1	1	0	1	93%
SR	93	Uvacsek et al. (2011)	0	1	0.5	1	1	1	0	NA	0	1	1	NA	NA	1	0	0	1	61%
SR	94	Vajiala et al. (2010)	0	0	1	1	1	0	0	NA	0	NA	0	NA	NA	0	0	0	0	23%
SR	95	Wagman et al. (1995)	0	0	0	1	1	1	1	NA	1	1	0	NA	NA	1	0.5	1	1	68%
SR	96	Wanjek et al. (2007)	0	1	1	1	1	1	1	NA	1	0	1	NA	1	1	1	1	1	87%
SR	97	Whitaker et al. (2014)	0	1	1	1	1	1	1	NA	1	0	0	NA	NA	1	0	0	1	64%
SR	98	Wroble et al. (2002)	1	0.5	1	1	1	1	1	NA	1	0	1	NA	NA	1	1	0	1	82%
SR	99	Yesalis (1988)	1	0	0	1	1	0	0	NA	1	NA	1	0	NA	1	1	1	1	64%
SR	100	Zenic et al. (2010)	0	0.5	0.5	1	1	0	0	NA	1	0	1	NA	NA	1	1	0	1	57%

QUAD-SAP Scores for Sample Analysis Prevalence

Study GP	REF#	Articles	1	2	3	4	5	6	7	8	9	10	Score
SA	112	Aguilar et al. (2017)	1	0	0.5	1	0	1	1	0	0	0	45.0%
SA	113	Aguilar-Navarro et al. (2019)	1	0	0.5	1	0	1	1	0	0	0	45.0%
SA	25	Al Ghobain (2017)	1	0.5	0.5	1	0	1	1	0	0	0	50.0%
SA	125	Aubel et al. (2019)	1	1	1	1	1	1	0	1	0	0	70.0%
SA	115	Bahr & Tjørnhom (1998)	1	0.5	0.5	1	0	1	1	0	0	0	50.0%
SA	122	Faiss et al. (2020)	1	1	1	0.5	1	1	1	1	0	1	85.0%
SA	123	Kuipers et al. (2007)	1	0.5	0.5	0.5	1	1	1	0	0	0	55.0%
SA	126	Maquirriain (2010)	1	0	0.5	1	0	1	1	0	0	0	45.0%
SA	116	Marchand et al (2017)	1	1	0	1	1	1	1	0	0	0	60.0%
SA	117	Mazzeo et al (2016)	1	0.5	1	1	0	1	1	0	0	0	55.0%
SA	118	Mercado et al. (2019)	1	0	0.5	1	0	1	1	0	0	0	45.0%
SA	119	Pereira & Sardella (2014)	1	0	0.5	1	0	1	1	0	0	0	45.0%
SA	80	Petroczi et al. (2011)	1	1	0.5	0.5	1	1	0	1	0	1	70.0%
SA	124	Sottas et al. (2011)	1	0.5	0.5	0.5	0	1	1	1	0	1	65.0%
SA	120	Van Eenoo & Delbeke (2003)	1	0	1	1	0	1	1	0	0	0	50.0%
SA	121	Vouillamoz et al. (2009)	1	1	1	1	1	1	1	1	0	0	80.0%
SA	127	Whitaker & Backhouse (2017)	1	1	1	1	1	1	1	1	0	0	80.0%