

Review

Not peer-reviewed version

Charting the Scientific Landscape of Indirect Estimation Models in Doping Prevalence Research: A Narrative Review with Bibliometric Analysis

Andrea Petróczi , [Dominic Sagoe](#) * , [Anna Kiss](#) , Sándor Soós , Razieh Chegeni , Annalena Veltmaat , [Maarten Cruyff](#) , [Peter van der Heijden](#) , [Olivier de Hon](#)

Posted Date: 23 March 2026

doi: 10.20944/preprints202603.1754.v1

Keywords: doping prevalence; sport, survey; indirect estimation; Randomized Response, narrative review; bibliometric mapping



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Review

Charting the Scientific Landscape of Indirect Estimation Models in Doping Prevalence Research: A Narrative Review with Bibliometric Analysis

Andrea Petróczi ¹, Dominic Sagoe ^{2,3,*}, Anna Kiss ^{4,5}, Sándor Soós ^{4,5}, Razieh Chegeni ^{3,6}, Annalena Veltmaat ⁷, Maarten Cruyff ⁸, Peter van der Heijden ^{8,9} and Olivier de Hon ¹⁰

¹ Faculty of Health and Sport Sciences, Széchenyi István University, Győr, Hungary

² Department of Psychosocial Science, University of Bergen, Bergen, Norway

³ Human Enhancement and Body Image Lab (HEBI Lab), Addiction Research Group, University of Bergen, Bergen, Norway

⁴ Department of Science Policy and Scientometrics, Hungarian Academy of Sciences (MTA), Budapest, Hungary

⁵ Faculty of Education and Psychology, ELTE Eötvös Loránd University, Budapest, Hungary

⁶ PROMENTA Research Center, Department of Psychology, University of Oslo, Oslo, Norway

⁷ Department of Sport and Sports Science, TU Dortmund University, Dortmund, Germany

⁸ Faculty of Social Sciences, Utrecht University, Utrecht, Netherlands

⁹ Southampton Statistical Sciences Research Institute, University of Southampton, Southampton, UK

¹⁰ Doping Authority Netherlands, Capelle aan den IJssel, Netherlands

* Correspondence: Dominic.Sagoe@uib.no

Abstract

Interpreting doping prevalence estimates generated through indirect estimation models (IEM) remains challenging for sport policy and governance due to wide variation in reported rates and methodological complexity. Building on Sagoe et al. (2024), we combined a critical narrative review of methodological and epistemic developments with a bibliometric analysis of publication trends, citation patterns, and collaboration networks, using a convergent parallel mixed-methods design. Across 52 records published between 2002-2026, this study maps the scientific landscape of IEM-based doping prevalence research. Findings show that IEM-based prevalence research is methodologically sophisticated yet institutionally dispersed and largely Eurocentric, reflecting a field still consolidating its standards and disciplinary identity. Over time, the focus has shifted from reporting prevalence rates to methodological critique and reanalysis of existing datasets. Reported prevalence estimates, ranging from 0 to 57.1%, are highly sensitive to modelling assumptions about athlete behavior in complex survey environments. While this trend strengthens rigor, it also complicates evidence synthesis for policy actors and risks undermining trust in IEM-based estimates if poorly communicated. Anti-doping organizations and researchers should treat IEM-derived prevalence as bounded indicators rather than definitive rates and integrate prevalence evidence with contextual data for transparent policy and public communication.

Registration: PROSPERO (CRD42022373691)

Keywords: doping prevalence; sport, survey; indirect estimation; Randomized Response, narrative review; bibliometric mapping

1. Introduction

The prevalence of doping remains one of the most frequently cited yet least well-understood phenomena in sport measurement [1]. Despite decades of surveillance and testing, a persistent gap remains between what anti-doping systems can detect and what athletes may actually do, which is

one of the most enduring paradoxes in sport governance) [2]. Even in the context of extensive testing programs and surveillance infrastructures, direct assessment captures only a small fraction of actual doping behavior. Biological testing detects incidences rather than prevalence, while conventional self-reports are vulnerable to denial, fear of exposure, and social desirability bias. In a domain where concealment is structurally embedded, empirical observation is necessarily partial. It is therefore intuitively appealing to follow the adage that, if one wishes to understand behavior, one should simply ask those involved. In principle, self-reporting offers a direct route to insight into athletes' engagement with prohibited practices. In practice, however, eliciting honest self-admission of doping is not simple, or straightforward. The potential consequences of disclosure [3,4], combined with the strong social stigma attached to doping [5], mean that even surveys conducted under conditions of assured anonymity cannot be assumed to yield truthful responses. Behind every prevalence statistic lies both a practical and an ethical dilemma: how to measure behaviors that individuals are strongly motivated to conceal. To address this challenge, researchers studying doping (or other socially sensitive or transgressive behaviors) have increasingly turned to indirect survey approaches designed to reduce response bias and enable more candid reporting.

1.1. Estimating the Prevalence of Sensitive Behavior

Indirect estimation models (IEM) comprise a family of innovative survey techniques designed to protect both respondents and researchers by creating safe survey conditions that go beyond conventional anonymity [6–8]. By design, IEM obscure the link between an individual's response and the sensitive behavior being assessed. Even when a respondent's answer to a survey item is known, it is impossible to determine whether it represents an admission or a denial of doping, because the response options are intentionally masked through a randomizing or unrelated mechanism. This feature provides a crucial procedural and psychological safeguard for respondents and ensures that individual identification — and by extension, any form of sanction or prosecution — is impossible.

Some IEM variants, including the Forced Response (FR) by Boruch [9], Kuk's model [10], and the Unrelated Question Model (UQM) [11,12], achieve protection through a different masking mechanism; they obscure which question—sensitive or unrelated innocuous—is being answered (Table 1). In these designs, only a random subsample of respondents receives the sensitive question, while others respond to an unrelated item with a known probability distribution. In other models, such as the Crosswise Model (CM) [13] and the Single Sample Count (SSC) [14,15], respondents are never required to directly admit the undesirable behavior. Instead, their answer to the sensitive item is combined with an unrelated question, producing a composite response that fully conceals whether the respondent has engaged in an undesirable behavior (e.g., using prohibited means in sport). The key protective feature is that only the respondent knows which question is being answered, not the researcher. Together, these mechanisms make IEM uniquely suited for collecting valid data on socially sensitive or prohibited behaviors such as doping, while maintaining the ethical integrity of both participants and investigators [7,8].

Table 1. Comparison of IEM used in doping prevalence studies focusing on respondent experience and face validity.

Model family / example	How the question is experienced by respondents	How respondents' answers are protected	Face validity: how it feels like a doping survey	Forced affirmative response and its implications	Detecting survey-instruction noncompliance
Combined-response models (e.g. Crosswise)	Respondents answer the doping question together with a	Individual answers are concealed by combining	High: all respondents perceive that they are answering the	No forced "yes"; protection relies on	Requires two parallel versions and a randomly split sample

Model)	neutral question, reporting only whether the answers match	responses to the sensitive question and other unrelated non-sensitive questions	doping question	ambiguity of combined answers	
Randomized-response models (e.g. Forced Response, Kuk's design)	Respondents follow instructions that sometimes require answering the doping question and sometimes require a preset answer	Protection is achieved because for the researcher, forced and genuine "yes" answers to the sensitive question are indistinguishable	Moderate: not all respondents feel they meaningfully answered the doping question	Yes: some respondents must say "yes" regardless of behaviour, which may reduce comfort and increase noncompliance	Requires two parallel versions and a randomly split sample
Question-substitution models (e.g. Unrelated Question Model)	Respondents answer either the doping question or a harmless question, determined by chance	Researchers cannot identify who answered which question (the sensitive or the unrelated question)	Moderate: only part of the sample directly answers the doping question	No forced "yes"; protection depends on question substitution	Requires two parallel versions and a randomly split sample
Count-based models (e.g. Single Sample Count)	Respondents report how many statements apply, without specifying which ones	Individual responses remain fully concealed through aggregation of the responses of which the sensitive question is only one of many	High: respondents feel included, but the doping question is indirect	No forced "yes"; protection comes from lack of item-level disclosure	Does not require two parallel versions if the non-sensitive questions are set to known prevalences (e.g., distribution of birth dates)

Analytical approaches to identifying and adjusting for survey instruction noncompliance are, in principle, applicable across all IEM. In practice, however, most models require empirical manipulation at the data collection stage, typically through the administration of two parallel versions of the survey or additional experimental conditions. Such pre-planned design features are a prerequisite for formally testing the presence and magnitude of noncompliance and for enabling retrospective statistical adjustments. Without this level of methodological preparedness, noncompliance remains largely unobservable and must be implicitly absorbed into prevalence estimates. One exception to this is the SSC method, which—unlike the other indirect models summarized in Table 1—does not require parallel sampling frames or additional experimental conditions. Its structure allows for post hoc assessment of response irregularities using a single dataset, thereby reducing logistical burden at the data collection stage. This practical advantage, however, comes with its own interpretive trade-offs and wider confidence intervals [15].

1.1.1. Behavioural Aspects

From a regulatory perspective, it is important to recognise that differences between IEM are not limited to statistical properties but also extend to how surveys are experienced by respondents. One key consideration is face validity, understood here as whether a survey clearly appears to be about doping prevalence and whether respondents feel that their participation meaningfully contributes to answering that question. Models with high face validity ensure that all respondents perceive themselves as answering the doping question, which may support engagement and compliance, but can also increase perceived personal risk if protection mechanisms are not well understood. Forced-response variants introduce another issue and present a distinct trade-off. By design, some respondents are required to give an affirmative (“yes”) response irrespective of their true behaviour. This feature enhances anonymity because the researcher cannot distinguish between genuine and forced admissions. At the same time, being instructed to say “yes” to a sensitive and normatively charged question like doping may feel uncomfortable or ethically troubling for some participants [16]. This discomfort can manifest as partial noncompliance, refusal, or evasive responding, potentially affecting data quality in ways that are not always visible in the final prevalence estimate.

Across IEM, several regulatory implications apply regardless of specific design. Prevalence estimates derived from these methods can support broad situational awareness, such as identifying whether doping is plausibly present beyond negligible levels or monitoring changes within the same population over time. However, they cannot support absolute prevalence thresholds, individual attribution, or direct enforcement logic. Differences between estimates should not be interpreted straightforwardly as differences in underlying behaviour, as they may instead reflect variation in anonymity protection, respondent comfort, or model assumptions. The main regulatory risk, therefore, lies not in the use of indirect models per se, but in the misinterpretation of their outputs. When prevalence figures are detached from their methodological and experiential context, higher estimates may be read as unequivocal indicators of regulatory failure, while lower or revised estimates may be viewed with suspicion or attributed to political or institutional motives. Appreciating how face validity, forced responses, and respondent experience shape these estimates is essential for preventing overconfidence, selective citation, and policy decisions driven more by numerical visibility than by evidential nuance.

1.1.2. Protecting Both Sides

An often-overlooked implication of this design is that IEM-based methods also protect the researcher. In traditional self-report surveys, researchers collecting identifiable data on prohibited behaviour could, in theory, become aware of individual admissions of doping. For investigators who are also practitioners bound by the World Anti-Doping Agency (WADA) Code, such knowledge could create an ethical dilemma arising from conflicting duties of confidentiality toward respondents and obligations to report known dopers to regulatory bodies (e.g. WADA, national anti-doping agencies, or sport federations). Because the IEM structure prevents anyone, including the investigator, from knowing which respondents have admitted to doping, it removes this potential conflict of duty. Doping prevalence can only be estimated at the aggregate level as the proportion of athletes statistically inferred to have engaged in doping, without identifying individuals. Thus, IEM perform a dual protective function by promoting honest disclosure among participants while safeguarding researchers from moral and professional jeopardy. In IEM’s mutually protective framework, respondents can answer truthfully without fear of exposure while researchers can collect valid data without risking ethical or legal compromise. Thus, the appeal of this approach in doping prevalence research is evident. When applied to doping, IEM have consistently revealed prevalence rates far exceeding those indicated by positive test results [2], which is frequently interpreted as a more realistic view of the hidden dimensions of sport [17,18].

1.1.3. Limitations of IEM

Despite their advantages, IEM are not without limitations. By design, these models rely on specific probabilistic assumptions. In the FR model, the probability of answering the sensitive question is known, while in the UQM the probability of answering the unrelated question is predetermined. Similarly, in the SSC and CM, the probability of providing an affirmative answer to the unrelated question is known. IEM uses this information to estimate the proportion of affirmative responses (i.e., the admission rate) to the sensitive target question concerning doping or other undesirable behaviours. However, these estimates depend on full respondent compliance and honesty, which cannot always be assumed. Despite the protection afforded by IEM, research has shown that social desirability bias can still influence responses [19], leading some participants to provide self-protective answers regardless of the methodological safeguards. Moreover, the relative complexity of IEM survey instructions places cognitive and attentional demands on respondents. Variations in reading comprehension, understanding of randomization mechanisms, or willingness to follow instructions precisely can result in careless or random responding. Respondents may also engage in satisficing, that is speeding through the survey without full engagement, which can distort prevalence estimates.

Both intentional and unintentional response errors introduce bias into the estimated prevalence of the target behaviour, in some cases substantially. To address these issues, methodological innovations have been proposed, including cheating detector variants [20,21] and parallel-form designs [22,23] that allow for post hoc adjustments of prevalence estimates. These adjustments are typically based on statistical assumptions about the likely extent and nature of noncompliance with survey instructions [24,25]. Such refinements reflect ongoing efforts to balance methodological rigour, cognitive feasibility, and respondent trust in research on sensitive and transgressive behaviours.

1.2. *Estimated Doping Prevalence and Its Interpretation*

The most comprehensive synthesis of IEM-based doping prevalence research is the systematic review and meta-analysis by Sagoe et al. [26], which reviewed 49 outputs published between 2002 and 2024 and aggregated prevalence estimations from 33 studies, covering more than 43,000 athletes. Across multiple IEM, including the Unmatched Count Technique, FR, CM, and related designs, the pooled lifetime prevalence was estimated at 22.5% among competitive athletes (14.3% past-year) and 17.2% among recreational sport participants (10.3% past-year). However, these estimates are accompanied by substantial interpretive uncertainty, not least because definitions of doping were inconsistent across studies and survey instruction noncompliance averaging nearly 29% was infrequently measured or reported.

For non-expert audiences, including policymakers and sport administrators, navigating divergent prevalence estimates produced by different indirect models is particularly challenging. Without a clear understanding of how these figures are generated and what they can meaningfully represent, estimates risk being compared across incompatible methods or selectively mobilized to support predetermined narratives. Prevalence estimates derived from IEM are intrinsically model-dependent and contingent on a series of methodological assumptions, including the functioning of randomization devices and respondents' comprehension and compliance with complex survey instructions. Compared with conventional self-report surveys, IEM impose higher cognitive demands and thereby introducing additional sources of bias and uncertainty [18,27–29]. Despite these limitations, IEM-based estimates are often treated in public and policy discourse as definitive indicators of the scale of doping, frequently stripped of their methodological conditions. Media coverage tends to privilege the most striking figures, while subsequent methodological re-analyses or refinements may be framed as minimizing the problem rather than as legitimate scientific scrutiny.

1.3. Research Context and Aims

Although systematic reviews and meta-analyses are essential for summarizing prevalence ranges, they offer limited insight into the epistemic processes through which estimates are generated, circulated, and rendered authoritative. Combining narrative synthesis with bibliometric mapping adds critical analytical depth into prevalence estimates, showing that they are not neutral metrics but products of evolving methodological traditions, collaboration networks, and underlying assumptions. Bibliometric mapping identifies patterns of influence, visibility, and methodological dominance, while narrative interpretation situates these patterns within broader scientific and policy landscapes.

IEM have been increasingly used to estimate doping prevalence since the early 2000s, generating a substantial yet uneven body of evidence [26]. To clarify how these estimates are produced, interpreted, and embedded within the scientific field, the present study extends the systematic review and meta-analysis by Sagoe et al. [26] through a critical narrative review combined with bibliometric mapping. This integrated approach aims to enhance methodological transparency and policy relevance by examining the structural, collaborative, and conceptual development of IEM-based prevalence research.

2. Methods

2.1. Study Design

This study employed a mixed-method design following the convergent parallel model [30]. In this approach, quantitative and qualitative components are conducted concurrently and then merged to achieve a more comprehensive interpretation. The bibliometric component quantitatively mapped publication trends, citation impact, and collaboration networks within the field of IEM in doping prevalence research. The narrative component complemented this by qualitatively examining the conceptual evolution, methodological debates, and theoretical underpinnings of IEM use in this domain. The integration of these two strands provided both structural and interpretive insights into the scientific landscape of doping prevalence estimation.

2.1. Data

Bibliographic data were extracted from the systematic review and meta-analysis conducted by Sagoe et al. [26], updated with three studies [31–33]. All 52 outputs were qualitatively analysed, and a subset of publications indexed in the Web of Science (k = 26) and/or Scopus (k = 29) databases was eligible for bibliometric analysis (Table 2). The latest database literature check for update was conducted in January 2026.

Table 2. Outputs reporting doping prevalence estimates using IEM, listed in chronological order.

References	Publication language	Type of output	Indexed in WoS	Indexed in Scopus
Musch and Plessner [34]	English	conference abstract / unpublished manuscript	no	no
Plessner and Musch [35]	German	book chapter	no	no
Pitsch et al. [36]	German	magazine	no	no
Simon et al. [37]	English	academic journal	yes	yes
Pitsch et al. [38]	English	academic journal	no	yes
Pitsch et al. [39]	German	book chapter	no	no
Pitsch et al. [40]	German	magazine article	no	no

Pitsch et al. [41]	English	book chapter	no	no
Striegel et al. [42]	English	academic journal	yes	yes
Stamm et al. [43]	German	academic journal	no	no
Pitsch and Emrich [44]	English	academic journal	yes	yes
Striegel [45]	German	book chapter	no	no
Breuer and Hallmann [46]	German	monograph	no	no
Dietz et al. [47]	English	academic journal	yes	yes
James et al. [48]	English	academic journal	yes	yes
Nakhaee et al. [49]	English	academic journal	no	no
Pitsch et al. [50]	German	book chapter	no	no
ADA Serbia [51]	Serbian	research report	no	no
Stubbe et al. [52]	English	academic journal	yes	yes
Duiven and de Hon [53]	Dutch	research report	no	no
Backhouse et al. [54]	English	research report	no	no
Dietz et al. [55]	English	academic journal	yes	yes
Frenger et al. [56]	English	academic journal	yes	yes
Schröter et al. [57]	English	academic journal	yes	yes
Fincoeur and Pitsch [58]	Dutch	academic journal	no	no
Franke et al. [59]	German	academic journal	yes	yes
Elbe and Pitsch [60]	English	academic journal	no	yes
Pitsch [61]	English	book chapter	no	no
Ulrich et al. [62]	English	academic journal	yes	yes
Boardley et al. [63]	English	academic journal	yes	yes
Seifarth et al. [64]	English	academic journal	yes	yes
Heller et al. [65]	English	academic journal	yes	yes
Nilaweera et al. [66]	English	conference abstract	no	no
Balk and Dopeide. [67]	Dutch	research report	no	no
Hilken et al. [68]	English	academic journal	yes	yes
Heyes [69]	English	PhD thesis	no	no
Petróczi et al. [67]	English	academic journal	yes	yes
Pitsch [71]	English	academic journal	yes	yes
Reiber et al. [72]	English	academic journal	yes	yes
Sayed et al. [73]	English	academic journal	yes	yes
WADA [74]	English	research report	no	no
Abdulrazzaq and Tareq [75]	English	academic journal	no	no
Balk et al. [76]	English	academic journal	no	no
Christiansen et al. [77]	English	academic journal	yes	yes
Ulrich et al. [18]	English	academic journal	yes	yes
Cruyff et al. [78]	English	academic journal	yes	yes
Robach et al. [79]	English	academic journal	no	yes
Sayed et al. [80]	English	academic journal	yes	yes

Sayed et al. [81]	English	academic journal	yes	yes
Schu and Haller [33]	English	academic journal	yes	yes
Pitsch and Christiansen [31]	English	academic journal	yes	yes
Sayed et al. [32]	English	academic journal	yes	yes

2.3. Data Analysis

Data analysis comprised two components including a critical narrative review and bibliometric mapping, followed by an integrative synthesis of findings from both approaches.

2.3.1. Critical Narrative Reflection

First, a narrative review was conducted to provide conceptual and methodological context for the development of IEM in doping prevalence research. This component aimed to synthesise the theoretical rationale, model evolution, and methodological debates underpinning the use of IEM, thereby situating the bibliometric findings within the broader scientific and applied discourse. This review is based on the same body of literature identified in the companion systematic review and meta-analysis [26], updated and supplemented by additional methodological and conceptual papers that informed the historical and theoretical development of IEM. Each article was examined for its contribution to the conceptual understanding or methodological refinement of IEM in the context of sensitive or transgressive behaviour research, focusing on (1) the rationale for using IEM in doping studies, (2) variations in model implementation and interpretation, (3) common methodological challenges such as instruction compliance, and (4) emerging solutions, including model extensions and cheating-detection variants.

2.3.2. Bibliometric Analysis

Bibliometric mapping involved examining temporal trends in outputs, authors and authors' institutional affiliations, dominant outlets (journals) and fields where doping prevalence estimation studies were presented. Research fields and topics in WoS and Elsevier's SciVal were catalogued and analysed for dominant patterns. Academic impact was assessed via time-normalised citation recorded in Web of Science (WoS), as well as from Scopus' Field-Weighted Citation Impact (FWCI) and SciVal Topic Prominence, which is a composite indicator that ranks a research topic's momentum by combining recent citation counts, Scopus view counts, and the average CiteScore of the journals in which the topic's papers appear. Citation analysis in this paper retained the conventional bibliographic details found in traditional citation indices, and augmented them with additional contextual information, including the citation statement, its surrounding context, and the location of the citation within the citing article [82–84]. In addition to examining citations at the level of individual articles, we also analysed studies according to the type of citation. To enhance the model's explanatory power, two additional features were incorporated. Nodes (outputs) were classified into categories based on the specific IEM employed in each article, and edges (citation links) were categorised according to the role of the citation whereby the citation statements represented by the links. The taxonomy of these roles was simplified into four categories: (1) Method – where the cited article was used for methodological purposes only; (2) Multiple use – where the citation served several purposes (e.g., methodological reference and conceptual discussion); (3) Other – encompassing non-central functions such as brief mentions, and (4) Secondary data analysis. Additionally, author overlap between cited and citing papers was examined to account for self-citation and collaborative influence.

Latent community structures were explored with network maps. First, we constructed a network map of all authors associated with the included outputs. Clusters were identified using the Louvain method implemented in the igraph package for R [85]. To evaluate the extent to which the included

outputs form a coherent line of research, a *local citation network* model was applied. This model represents the network of citations among the included studies by considering both incoming and outgoing citations restricted to this set of papers. The background and interconnectedness of the selected papers in terms of *research communities*, were evaluated in co-document network based on shared authorships. This model is conceptually the inverse of a conventional co-author network meaning that rather than connecting authors who have written together, it connects papers that share one or more authors. In this framework, two papers (A and B) are linked if they have at least one common author. By focusing on publications rather than individuals, the co-document network captures the intellectual structure of the field through patterns of shared authorship. This approach is particularly useful for identifying research communities or “intellectual camps,” assessing the cohesiveness of the literature, and detecting potential bridging papers that connect otherwise separate groups. In the present study, the co-document network was applied to explore how studies employing IEM in doping research cluster around shared expertise, methodological preferences, and research focus (topics). We applied the Louvain algorithm to detect coherent subgraphs representing clusters of closely related publications [86].

Network visualisation and interpretation were carried out to aid the qualitative understanding of the bibliometric structures and to facilitate the interpretation of how methodological preferences, collaborative patterns, and intellectual lineages structure the field of doping prevalence estimation with IEM. The resulting co-authorship network, and citation network were visualised using Cytoscape web 1.0.5 (www.cytoscape.org). Co-document networks were visualised using force-directed layouts, which position nodes based on the strength and density of their connections.

2.3.3. Assessment of Overall Evidentiary Strength

Overall evidentiary strength was assessed using a modified version of the framework proposed by Palmateer et al. (p. 846) [87], adapted for IEM-based doping prevalence research. Prevalence estimates were grouped into 5% bins and cross-tabulated by IEM model. Cells recorded the number of studies and were coded by adjustment for survey instruction noncompliance (adjusted vs. unadjusted) and by analytical status (primary vs. secondary). Cumulative evidence was qualitatively interpreted as sufficient, tentative, insufficient, or none, based on the convergence and robustness of available primary studies.

2.3.4. Data Integration

Following the principles of the convergent parallel design [30], the bibliometric, narrative, and evidentiary assessment components were conducted and analysed independently, with results integrated during the interpretation phase. Quantitative findings from the bibliometric mapping, such as publication trend over time, outlets, research topic classifications, citation structures and collaboration networks, were compared and cross-referenced with qualitative insights from the narrative synthesis, including theoretical debates, methodological adaptations, and conceptual developments. Integration was achieved through interpretive triangulation, allowing complementary strands of evidence to inform each other.

3. Results

The number of published studies increases only gradually over the observed period, indicating a relatively slow expansion of the evidence base, with the number of outputs fluctuating between two and five per year (see Figure 1). In contrast, the number of unique researchers involved shows a more pronounced upward trend. This divergence suggests that, while growth in outputs remains modest, IEM-based doping prevalence estimation is attracting an increasingly broader research community, pointing to a slowly rising methodological interest and collaborative engagement beyond what is reflected by publication counts alone.

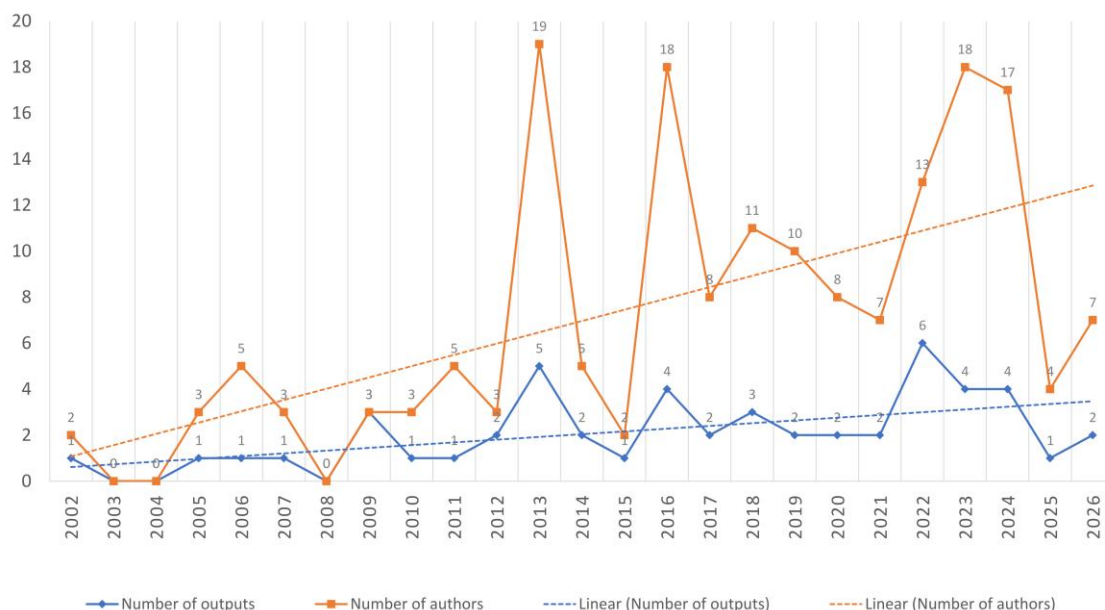


Figure 1. Temporal trends in the use of IEM for estimating doping prevalence in the number of published outputs and contributing authors.

3.1. Publication Patterns

Among the 52 records, outputs were dominantly research articles ($k = 33$), followed by book chapters and monographs ($k = 7$), publicly available reports ($k = 3$), unpublished research report ($k = 2$), published conference abstracts ($k = 2$), magazine articles ($k = 3$), PhD thesis ($k = 1$), and unpublished manuscript ($k = 1$). Adoption and early-year-applications of IEM to estimate doping prevalence was dominated by authors from Germany (see Figure 2) since 2002. The only other countries where researchers demonstrated sustained involvement in study with IEM were the UK and Netherlands in both national and international collaborations.

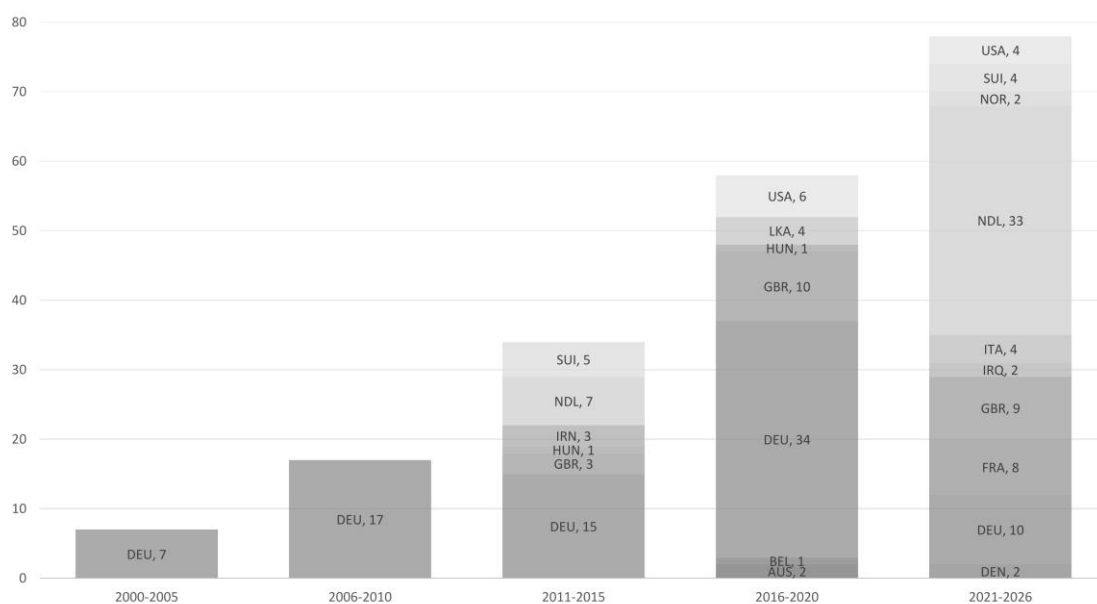


Figure 2. Trends in the use of IEM for doping-prevalence research by author diversity across countries.

Most outputs were published in English ($k = 39$), followed by German ($k = 9$), Dutch ($k = 3$) and Serbian ($k = 1$), with some overlap and duplication between English- and German-, and English- and Dutch language versions. These instances represent duplicate publications, where identical datasets and results were disseminated across multiple outputs. Specifically, two German-language studies [39,40] reported the same data later published in English by Pitsch and Emrich [44], and also included material from an earlier investigation [36] that was subsequently re-presented in another publication [38]. Data from an unpublished manuscript [34], included here with permission, were presented as a conference abstract, and published in German as a book chapter [35]. Similarly, results from a recent Dutch doping prevalence study [53,67] were republished in Balk et al. [76]. Pitsch [71] and Christiansen et al. [77] reported identical data, with further subgroup analysis presented in Pitsch and Christiansen [31].

A second category comprised secondary analyses, where data were re-analysed using refined algorithms or alternative assumptions. Ten studies in three sets fell into this category. Ulrich et al. [62] and Petróczi et al. [67] were conducted in the same setting and shared one dataset but applied different IEM variants. Reiber et al. [72] and Ulrich et al. [18] subsequently re-analysed the same data generated with UQM [62] and SSC [67] models respectively, testing a different hypothesis and introducing revised assumptions about the magnitude and causes of noncompliance with survey instructions. Likewise, data first reported by Cruyff et al. [78] where two sets of results were presented including one set for the unadjusted prevalence estimation (assuming full compliance with survey instructions) and a set adjusted for self-protective responding. These were later re-examined in Sayed et al. [32] to assess the potential impact of random responding (i.e., participants accelerating through the survey by selecting responses at random). The third set revolved around the Kuk's model and comprised two parent studies [53,68] and two subsequent re-analysis to investigate the impact of timeframe reference (i.e., lifetime (ever) and current (last year)), and evasive responding [79,80].

3.2. Publication Channels and Research fields

The selected studies are distributed across a wide range of publication outlets and research fields (Tables S1 and S2), reflecting substantial dispersion despite a shared substantive focus. Of the 33 journal outputs identified (32 with DOIs; 26 indexed in Web of Science), 12 were concentrated in three journals (*PLOS One*, *Sports Medicine/Sports Medicine–Open*, and *Performance Enhancement & Health*), while the remaining 22 appeared in 22 different outlets. Although 18 of the 26 journals were ranked in the top quartile (Q1) of the Scimago Journal Ranking, this dispersion suggests that IEM-based doping prevalence research has been evaluated by diverse peer-review communities with varying levels of methodological expertise. Notably, only about half of the journals targeted a sport science readership, indicating that doping prevalence often serves as a test case for methodological development rather than the sole focus of inquiry.

Disciplinary clustering aligns with this pattern. Sport science and psychology journals tend to prioritise applied prevalence estimates, whereas statistical and methodological journals focus on model validation and analytical refinement, reinforcing the interdisciplinary yet fragmented nature of the field. WoS subject categorisation further amplifies this dispersion. Although all 26 empirical studies examined doping prevalence in sport, they are indexed across 34 subject categories in WoS (Figure 3), giving the appearance of a broad evidence base despite substantial overlap in data, models, and author networks.



Figure 3. Word cloud of Web of Science subject categories, where font size reflects category prominence (i.e., the more frequently a study is assigned to a given subject area, the larger its label appears).

At its core, the field is anchored in psychology, sport sciences, and public, environmental and occupational health, framing doping primarily as a sport-related behavioural health issue. Additional classifications in psychiatry and substance abuse further accentuate a clinical framing, despite limited engagement with diagnosis or treatment. Methodologically driven categories such as mathematics and mathematical methods in the social sciences contribute disproportionately to the field's visibility. Social science perspectives remain uneven, with sociology moderately represented and governance- or policy-oriented fields largely marginal. Output-level categorisation is given in Supplementary material (Table S3).

Looking from a different angle, the outputs were distributed across six high momentum SciVal Topics, with the largest concentrations found in *doping policies and athlete integrity in sports* ($k = 11/30$) and *randomized response techniques for sensitive surveys* ($k = 11/30$), reflecting both the centrality of doping related concerns and the methodological advances used to estimate their prevalence. The remaining four topics comprise *research on the health risks of anabolic steroid use* ($k = 3$), *prescription drug misuse and cognitive enhancement* ($k = 3$), *nutritional supplement use and performance* ($k = 1$), and *erythropoietin related doping and detection methods* ($k = 1$).

3.3. Framing of Doping in Titles and Publication Contexts

An analysis of the publication titles and journal outlets reveals clear patterns in how doping is conceptually framed across disciplines. These patterns mirror the disciplinary homes of the journals, highlighting how scientific communities construct the meaning and boundaries of doping and doping prevalence. Specifically, titles published in sport science and medicine journals (e.g., *Sports Medicine*, *Scandinavian Journal of Science & Medicine in Sport*, *Journal of Sport Sciences*) typically adopt an epidemiological and empirical framing, positioning doping as a measurable phenomenon. Terms frequently used in titles emphasise quantification, method, and comparability with terms such as prevalence, estimation, frequency, and use. This reflects a biomedical and sport science discourse, where doping is treated as a population-level health and/or integrity issue requiring methodological rigour and large-scale evidence.

In contrast, publications in journals such as *Addiction*, *Drug and Alcohol Dependence*, *Performance Enhancement & Health*, and *Psychology of Sport and Exercise* frame doping as a behavioural or psychosocial phenomenon. Here, the lexical field shifts from prevalence to use, often in combination

with *attitude, susceptibility, or vulnerability*, which suggests an interpretive stance oriented toward individual human problem behaviour rather than population measurement. A third cluster, comprising journals such as *the International Review for the Sociology of Sport, Journal of Criminal Law, Criminology and Criminal Justice*, adopts a moral, regulatory, or sociological framing. In these, doping appears as a social deviance or policy problem, embedded in wider issues of governance, integrity, and the health of elite sport systems. The recurring use of terms like *risk management* and *sport-induced substance use* reflects this more normative and institutional perspective.

3.4. Evidentiary Synthesis

The synthesis of IEM-based prevalence estimates reveals substantial heterogeneity across methods, samples, and analytical approaches (Table 3, Table S4). To assess the strength of evidence within this diverse body of work, we consider the density of estimates falling into specific prevalence 'bins'. The number of evidence points exceeds the number of unique studies because many publications report multiple estimates across subgroups, time points, or analytical specifications.

Table 3. Evidence strength mapping across the IEM-based prevalence studies.

	0–5	6–10	11–15	16–20	21–25	26–30	31–35	36–40	41–45	46–50	51–55	56–60	61–65	66–70	71–75	76–80	80+	
F	1	6	5	1	2	2	1											
R	4																	
U	3	9	5	1	1		1		1			2						
Q																		
M																		
C	1		1	1	3	3	3	2	1		1	1	1				1	
M																		
SS		7	1	2	1	4	1		4	1								
C																		
K	6	5	3	5	1													
uk																		
's																		

Blue: unadjusted; Orange: adjusted for noncompliance (re-analysis with a set of assumptions about noncompliance) and subgroup analyses. Shading indicates frequency of studies in that particular cell (darker shades denote higher frequency counts). Limited to peer-reviewed outputs (journal articles and book chapters). FR: Forced Response, UQM: Unrelated Question Model, CM: Crosswise Model, SSC: Single Sample Count.

The overall picture in Table 3 shows that evidence is unevenly distributed across IEM families, reflecting shifts in methodological popularity over the past 25 years. While prevalence estimates span a wide range, consistent with variation in athlete populations and definitional differences, the strongest and most consistent concentration lies within the lower prevalence bins. Across methods, designs, and operationalisations, repeated clustering in the 0–20% range indicates a more stable and coherent empirical signal in this part of the distribution. In contrast, higher prevalence estimates appear less frequently and are more closely tied to specific methods, analytical assumptions or a unique sample. This overall picture appears to be congruent with the more nuanced meta-analytical synthesis presented in Sagoe et al. [26], indicating sufficient evidence up to 25%. Higher prevalence estimates off near or above 50% appear to be inconclusive or derived from a single study with small sample.

The overall picture in Table 3 shows that evidence is unevenly distributed across IEM families, reflecting shifts in methodological popularity over the past 25 years. While prevalence estimates span a wide range, consistent with variation in athlete populations and definitional differences, the strongest and most consistent concentration lies within the lower prevalence bins. Across methods, designs, and operationalisations, repeated clustering in the 0–20% range indicates a more stable and coherent empirical signal in this part of the distribution. In contrast, higher prevalence estimates appear less frequently and are more closely tied to specific methods, analytical assumptions or a unique sample. This overall picture appears to be congruent with the more nuanced meta-analytical synthesis presented in Sagoe et al. [26], indicating sufficient evidence up to 25%. Higher prevalence estimates off near or above 50% appear to be inconclusive or derived from a single study with small sample.

3.5. Scientific Impact

The average Field-Weighted Citation Impact (FWCI) of the included studies was 1.771 ($SD = 2.429$; median: 0.905; range: 0.00–10.38). Overall, the scientific impact of the corpus is above the international average, as median value reaches and the mean value exceed the global field-normalised benchmark of 1.0. The slightly lower median reflects a skewed distribution driven by a small number of highly cited outliers.

A more nuanced picture emerges when examining the temporal distribution of citation scores. Despite the intrinsic age normalisation of the MNCS metric, maintaining a three-year citation window remains advisable for reliable impact assessment. As illustrated in Figure 4, most outputs published between 2017 and 2025 cluster around or above the world average (1.0; black dotted line). Excluding the top and bottom 10 percent, the trimmed mean yields a still relatively high mean citation score of 1.53 (green dotted line), indicating that the field's influence has been both sustained and robust over time.

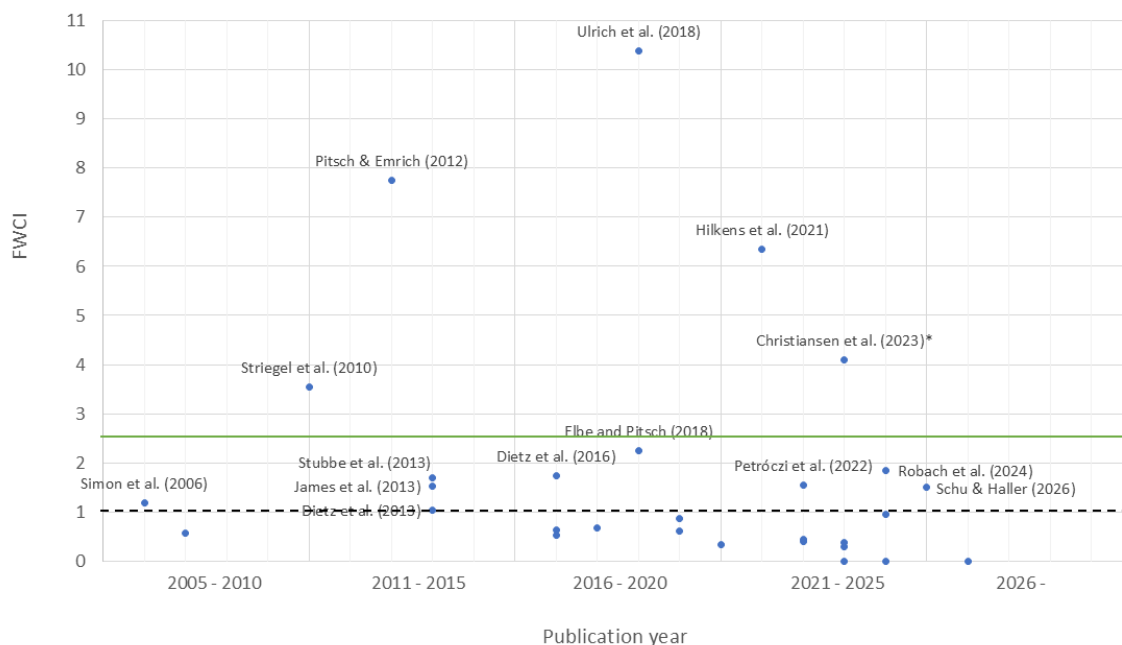


Figure 4. Scientific impact of the included studies, represented by their Field-Weighted Citation Index (FWCI) scores. Studies with $FWCI < 1$ are indicated but not individually labelled to preserve graph readability.

Based on the Scopus data (Table S3), the dataset demonstrates sustained topical relevance and strong, albeit heterogeneous, citation performance across the major strands of anti-doping research. Field-Weighted Citation Impact (FWCI) within these topics shows considerable variability, although several publications, particularly those within the most prominent topics, exceed global citation.

Topic-level impact scores with a mean FWCI of 2.00 ($SD = 2.71$) for *Randomized Response Techniques for Sensitive Surveys* and 1.77 ($SD = 2.47$) for *Doping Policies and Athlete Integrity in Sports* suggest that a field is equally split between method development and its application to assess the prevalence of doping in sport. The remaining four topics comprise research on the health risks of anabolic(-androgenic) steroid use ($k = 3$, mean FWCI = 2.17, $SD = 2.46$), prescription drug misuse and cognitive enhancement ($k = 3$, mean FWCI = 1.99, $SD = 2.77$), nutritional supplement use and performance ($k = 1$, FWCI = 1.75), and erythropoietin-related doping and detection methods ($k = 1$, FWCI = 1.85).

3.6. Authors and Authorship

One-hundred unique authors contributed to the literature on doping prevalence estimation with IEM, appearing 197 times (see Table S5). Among them, only 29 authors (29.0%) contributed more than one output, and only 12 authors (12.9%) had three outputs or more: *Pitsch* ($k = 13$), *Cruyff* ($k = 10$), *Petróczi* ($k = 10$), *Simon* ($k = 9$), *Ulrich* ($k = 8$), *Emrich* ($k = 7$), *Dietz* ($k = 7$), *Van der Heijden* ($k = 8$), *Sayed* ($k = 7$), *De Hon* ($k = 5$), *Striegel* ($k = 5$) and *Frenger* ($k = 3$). The collaboration pattern among the authors who have worked with IEM to estimate doping prevalence, visible in co-authorships, offers an intriguing picture (Figure 5). Authors in the corpus formed two unconnected clusters of different sizes, and six unconnected research groups. The small cluster is a tightly knit group of multiple jointly authored outputs centred around *Pitsch*. The large cluster is an amalgamation of three loosely connected groups around *Ulrich* and *Cruyff*, with *Petróczi* serving as a bridge between the other two.

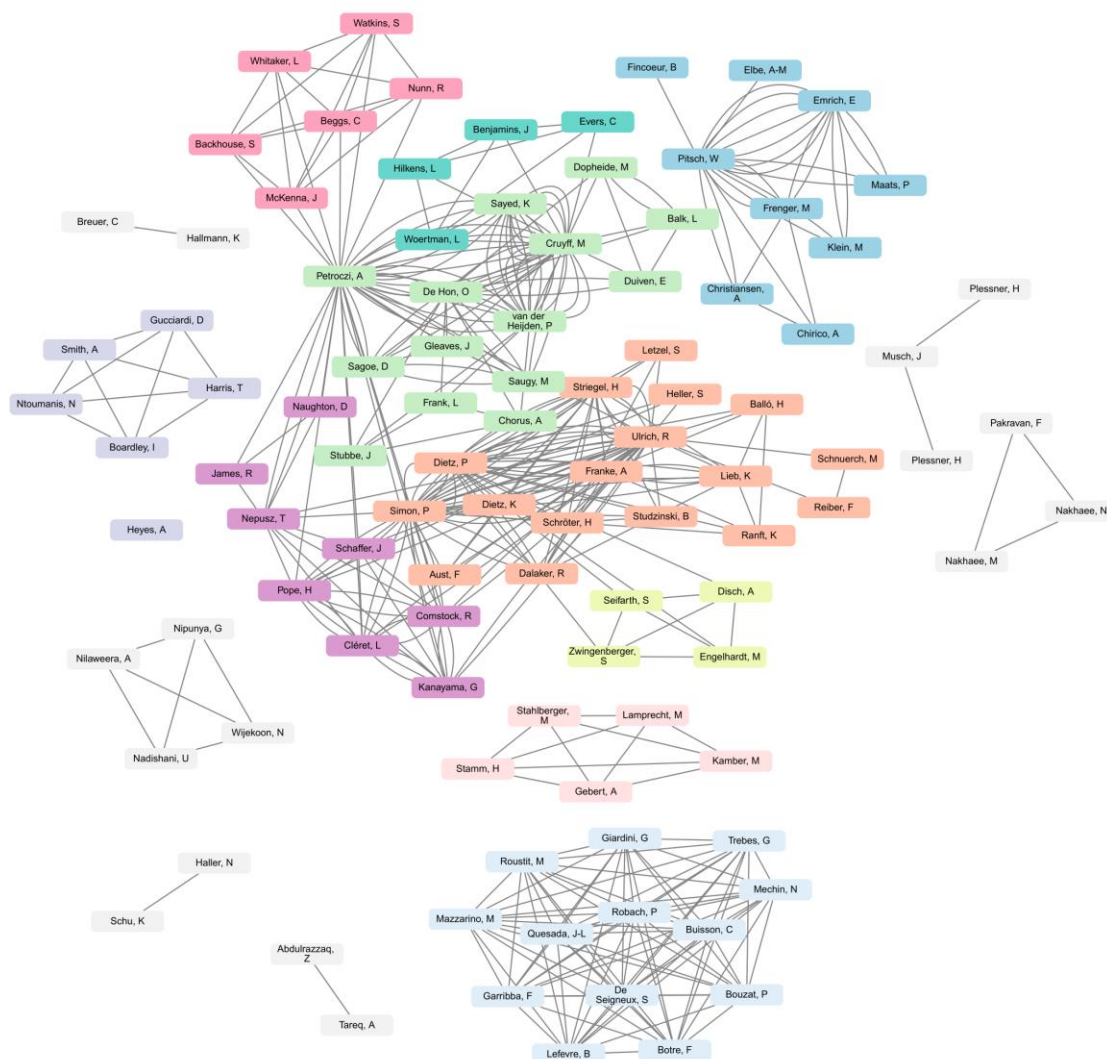


Figure 5. Co-authorship collaboration network among the included studies, with colours indicating empirically identified clusters.

Across the corpus, men were the majority among those developing, refining, or applying IEM to doping prevalence, accounting for 71% of all authors. Gender imbalance was even more pronounced among lead contributors. Of the 48 authored outputs, 39 (81.2%) listed a male first author, and among the 44 outputs where a corresponding author could be identified (some research reports did not specify one), only 14 (31.8%) were women. Last authorship was not analysed due to varying disciplinary conventions in author order across the contributing fields. This observed gender imbalance is not merely an equality statistic but may carry epistemic implications for how research questions are framed and which methodological approaches are privileged [88].

3.7. Research Communities

The co-document network based on shared authorship reveals the underlying structure of research communities contributing to IEM-based prevalence estimations applied to doping. The network, presented in Figure 6 displays a clear community organization among the included outputs. For ease of interpretation, detected communities (sets of outputs) are colour-coded by clusters.

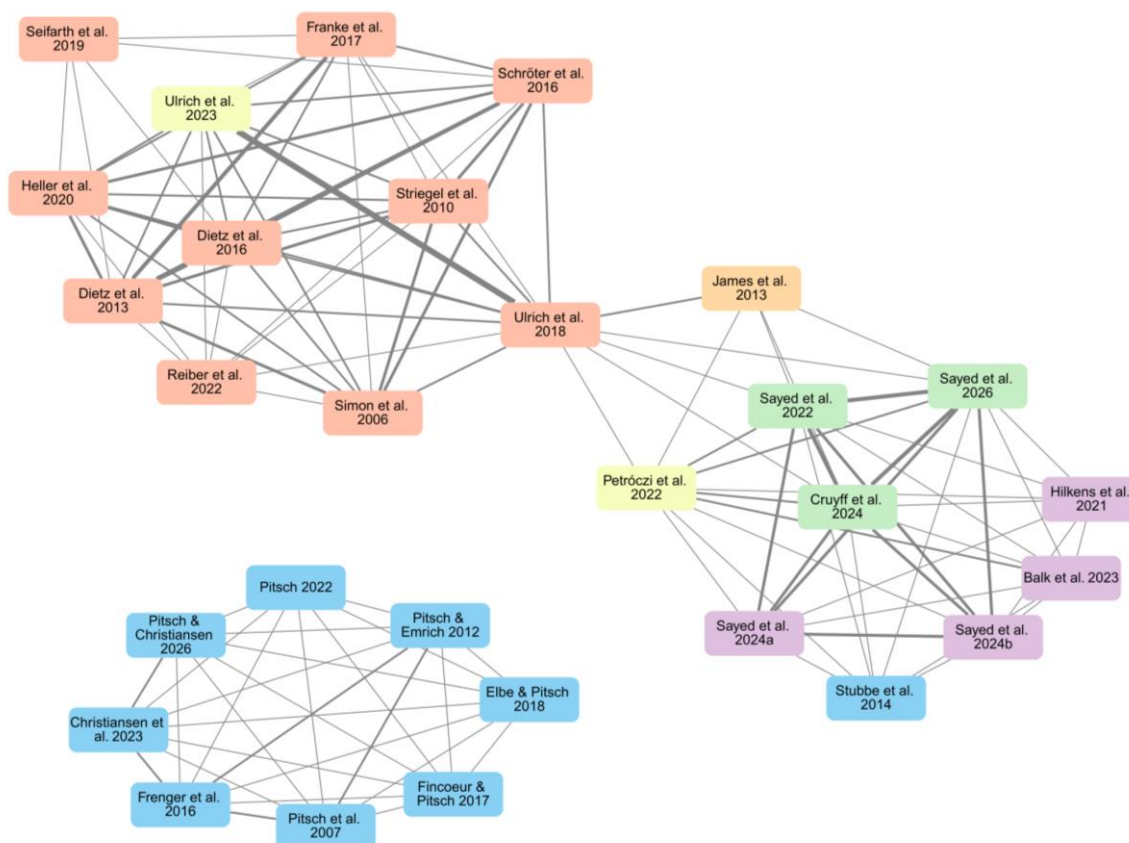


Figure 6. Co-document network colour-coded by the IEM used (blue: FR, green: ECWM, orange: SSC and UQM; pink: UQM; purple: Kuk's model; yellow: SSC); line thickness reflects the number of shared authors, with thicker lines indicating greater overlap between author groups.

As Figure 6 shows, two distinct and unconnected components emerged. The smaller component represents a fully connected group centred around *Pitsch*, indicating a tightly knit collaboration network with limited external connections. The larger component is more complex, comprising two coherent but only loosely interconnected subgroups around *Ulrich* and *Cruyff*. These subgroups are linked through a bridging publication —and more precisely, through *Petróczi*—whose authorship

on Ulrich et al. [62] connects the two otherwise separate clusters. This structural pattern mirrors the configuration observed in the overall author collaboration network (see Figure 5), suggesting that within-sample citation may be influenced more by existing collaborations and self-citation than by direct engagement with external scientific content. Notably, each community spans multiple publication years, indicating stable and sustained collaboration over time rather than short-term or project-specific partnerships. It can also be observed that certain author groups display a consistent preference for specific IEM variants. This pattern suggests that the selection of a model may not be determined purely by rational or technical considerations—such as selecting the most appropriate tool for a given research question or population—but is also shaped by familiarity, available expertise within the research team, and the legacy of prior collaborations. Language, training background, and beliefs about the truth (i.e., what is *true* prevalence) and best or most valid model appear to reinforce these preferences.

Overlaying Scopus SciVal Topics onto the co-document network provides a more nuanced view of the literature landscape. Just as in Figure 6 showing fairly coherent overlap between clusters and IEM employed, two dominant clusters for SciVal topics also emerged (Figure 7). Of these, one centred on doping-prevalence research within the topic *Doping Policies and Athlete Integrity in Sports*, and largely linked to the work of Pitsch and colleagues. The other cluster focused on methodological innovation within *Randomized Response Techniques for Sensitive Surveys*, with key contributions from Cruyff, Sayed, and Petróczi. The remaining publications and authors are distributed across four additional topics, reflecting the diversity of research directions within the field.

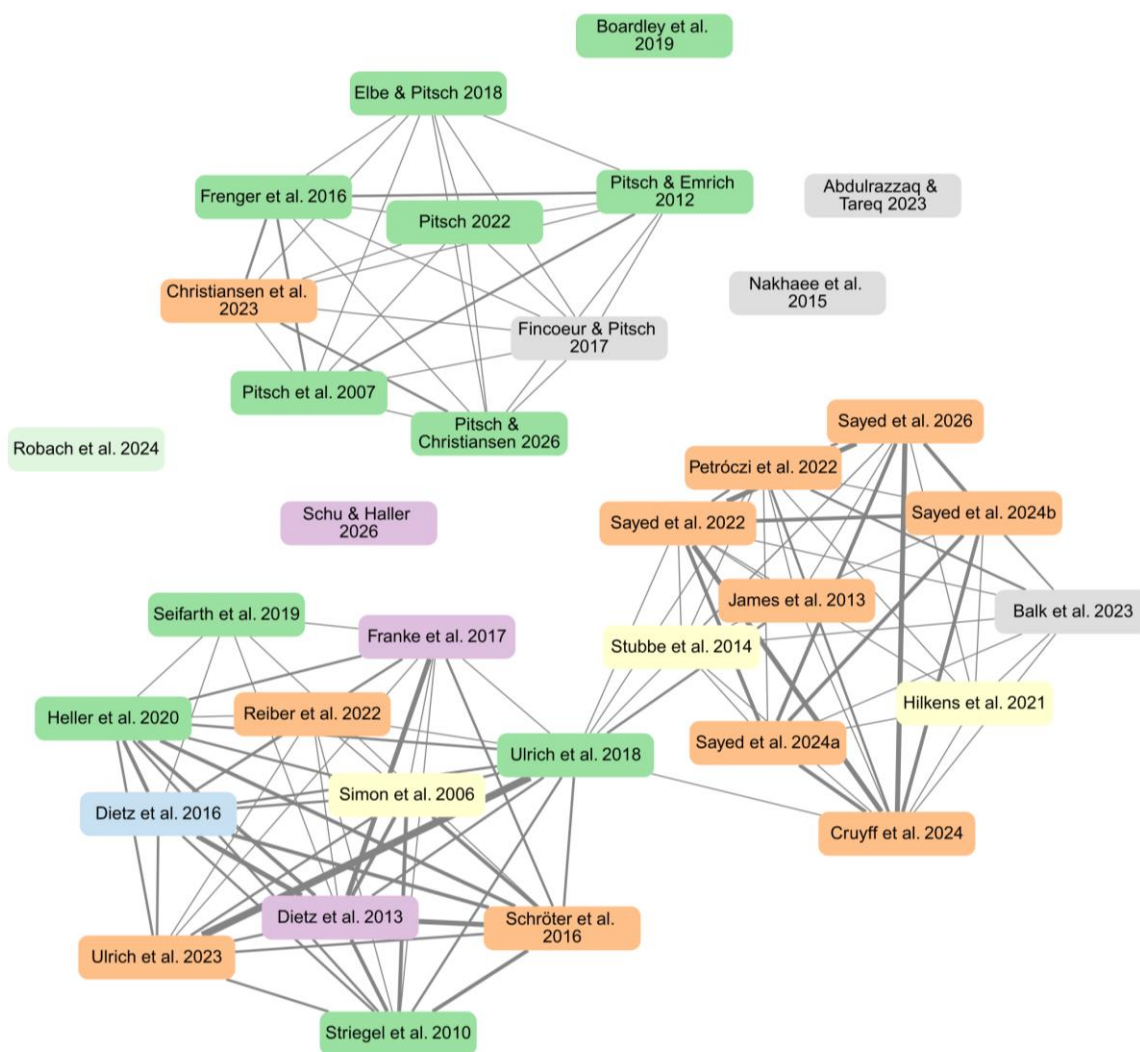


Figure 7. Co-document network colour-coded according to Scopus SciVal topics (green: *Doping Policies and Athlete Integrity in Sports*; orange: *randomized response techniques for sensitive surveys*; purple: *prescription drug*

misuse and cognitive enhancement, yellow: *research on the health risks of anabolic steroid use*; pale green: *erythropoietin related doping and detection methods*; blue: *nutritional supplement use and performance*.

This positional mapping is important because it highlights the conceptual role of each body of work. Studies primarily concerned with methodological development, refinement, and validation often generate prevalence estimates as secondary outputs of their analyses. These values should therefore be interpreted with caution. Using isolated or selectively extracted figures from method-focused studies as direct evidence of doping prevalence risks misrepresenting the intended scope and limitations of the research.

3.8. Local Citation Network

The local citation network model was used to evaluate the extent to which the included studies form a coherent line of research (Figure 8). The most informative feature of the network is that it comprises a single connected component, indicating that all articles in the sample are directly or indirectly linked through citation relations, with no isolated nodes. Beyond this overall connectedness, network-level measures show that distances within the graph are small and the network is relatively compact overall (average shortest-path length = 2.16), suggesting that ideas and methods diffuse across this literature in just a few intermediary citations, with early method papers acting as hubs and recent statistical developments forming the far end of the knowledge chain. The diameter (d) of the network is four, meaning that the longest shortest path between any two papers involves only four citation links (*Simon et al. 2006* → *Striegel et al. 2010* → *Ulrich et al. 2018* → *Sayed et al. 2022* → *Cruyff et al. 2024*). Together, these structural features describe a highly cohesive research line in which successive studies display continuous awareness of prior work in the field.

Several core papers, such as Pitsch et al. [38], Striegel et al. [42], Ulrich et al. [62] and Dietz et al. [55], serve as key reference points for later outputs, forming the backbone of the citation network. To further explore the *semantics of citation flow* beyond structure, we refined the local citation network by incorporating the type of citation relationship. Node colours represent the IEM model applied for prevalence estimation, while edge colours denote citation types. The taxonomy of citation roles was simplified into three categories: method, multiple use, and other mentions. Although these categories differ conceptually, the vast majority of citations fell under the mentioning type (acknowledging another study without direct relevance), while substantive method and multiple use citations were treated as indicators of knowledge transfer.

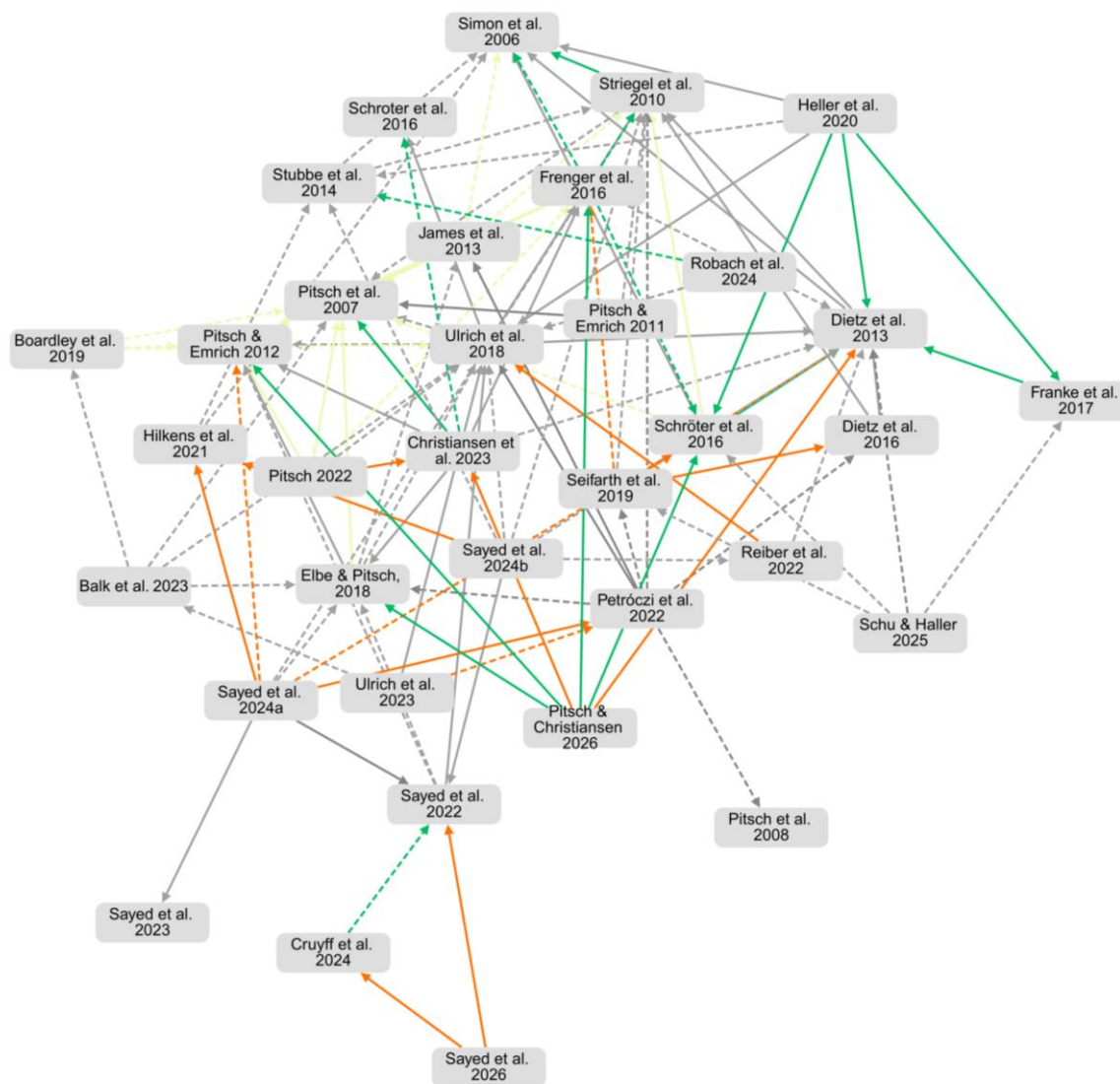


Figure 8. Within-corpus citation network, with edge colours denoting citation type (green = method; yellow = multiple use; grey = other; orange = data/secondary analysis). Solid lines indicate shared authorship between citing and cited outputs; dashed lines indicate no overlapping authorship.

The network depicted in Figure 8 exhibits both structural cohesion and functional connectivity. The citation flow consists predominantly of strong links, indicating that individual studies tend to engage with the methods, findings, or assumptions of preceding work. Importantly, these strong links do not necessarily imply endorsement because critiques and refinements can also generate dense citation connections. Based on author patterns, strong links within clusters are more likely to reflect methodological continuation and knowledge transfer, whereas links between clusters represent critical comparison or methodological debate.

When incorporating information about the specific IEM employed, an even more granular picture emerges. Two models dominate the corpus: the FR and UQM. The most frequent citation connections occur within these same-model pairings (FR → FR; UQM → UQM). However, cross-model links are also common, suggesting a degree of methodological awareness and continuity across the research community. Given the technical complexity of IEM, the emergence of entirely new research groups without prior collaboration or co-authorship is rare, and when it happens, it tends to be a one-off research enterprise.

Secondary data analysis occurred only when researchers conducting the re-analysis had been involved, in some capacity, in the original data collection or primary analysis. Such involvement was recognisable either through overlapping authorship indicating shared research communities across

studies or through contributions made via commissioned work, which may not appear in the citation network but were explicitly acknowledged in the publications. Across the corpus, no instances were identified in which an entirely independent research team re-used or re-analysed data generated by others.

3.9. Network cohesion, weak ties, and brokerage

Nearly half of all citation links exhibit overlapping authorship. Nearly half of the 110 links (49, 44.5%) involved at least one author appearing on both the citing and cited papers. Within the largest component, two prominent subgroups are visible: a densely connected cluster centred on *Striegel* and *Dietz* (frequently linked with *Ulrich*), and a looser constellation around *Petróczi*, *Sayed*, and *Stubbe*. This pattern consistent with a small expert base and repeated team-level collaborations. These self-referential patterns reinforce the presence of *invisible colleges* - informal, cohesive communication circles that organize knowledge flows within specialties and shape the growth of research areas [89,90].

From a network-theoretic perspective, the configuration we observe, namely dense internal linkages with selective cross-cluster connectors, aligns with classic theories of diffusion via weak ties [91] and brokerage across structural holes [92,93]. In such structures, a small number of bridges (e.g., citations linking Striegel et al. [42] to Ulrich et al. [62] and to Petróczi et al. [67] carry ideas between otherwise segregated subgroups, enabling cross-fertilisation that dense intra-cluster ties alone cannot deliver. Taken together, tight intra-cluster linkages appear to enhance conceptual coherence and speed method transfer within the two main camps, yet their concentration within closely connected teams also risks insularity, potentially limiting cross-paradigmatic exchange and slowing broader theoretical integration, which is a trade-off long noted in the literatures on invisible colleges, weak ties, and structural holes [89–92,94].

4. Integrated Results and Narrative Insights

Against the rich literature on IEM spanning over half a century [7,8,95], their application to doping only began around the turn of the millennium [35,36,38–40], with the first full publication in English appearing in 2006 [37]. Our findings indicate limited variability in study origin, with the majority of studies included in the meta-analysis conducted in European countries. Bibliometric analyses revealed that this trend was primarily driven by the dominance of two closely linked but distinct research groups in Germany. Over time, however, the trends show the emergence of new research groups in the United Kingdom and the Netherlands. WADA's establishment of a Working Group on Prevalence of Doping in Sport (2017–2023), with its focus on survey development (WADA, 2022), also facilitated the observed expansion in outputs, authorship, and diversity of IEM applications. Preferences for specific models are notable. For instance, the research group led by Ulrich predominantly applies the UQM, while Pitsch and colleagues favour the FR model. The CM and its variants have commended recognition in the field since its adoption by the WADA's working group, leading to a series of field testing [74,95] and methodological refinements [32,73,78] over recent years.

Indirect estimations of doping prevalence is a research field that is methodologically innovative yet structurally fragmented, with important implications for how doping prevalence estimates are produced, circulated, and interpreted. Combining quantitative indicators (e.g., publication patterns, outlets, and temporal trends) with qualitative analysis of research aims and framing provides a multidimensional picture of the intellectual development and epistemic orientation of IEM-based doping prevalence research. Across its development, the thematic focus of the field has shifted markedly. Early studies (2006–2012) were primarily concerned with demonstrating the feasibility of indirect methods, most notably RRT, for estimating hidden doping behaviour in elite and fitness sport contexts. These contributions were typically framed as proof-of-concept studies, aimed at showing that indirect questioning could yield plausible prevalence estimates where direct

approaches failed. During the subsequent period (2013–2019), the field expanded both empirically and conceptually.

Researchers increasingly embedded prevalence estimation within broader behavioural frameworks, examining gateway hypotheses, cognitive doping, and supplement use, while also extending empirical attention beyond elite sport to recreational and sub-elite populations. This phase was characterised by greater methodological experimentation, including the comparative application of multiple indirect techniques within the same samples. From around 2020 onwards, a pronounced methodological turn is evident. Recent studies increasingly focus on the development, critique, and refinement of IEM themselves, with explicit attention to sources of bias, evasive responding, instruction noncompliance, and potential inflation effects. This has been accompanied by the re-analysis and re-interpretation of earlier, high-profile prevalence estimates in light of new empirical and analytical insights. Collectively, these developments signal a shift in the field's core question from *"how prevalent is doping?"* toward *"how trustworthy and interpretable are our estimates?"*, and with the latter giving way to method-driven, nuanced re-analyses that take noncompliance into account for improved validity

Bibliometric patterns in publication outlets further reinforce this interpretation. The literature is highly dispersed across journals, but clusters around four broad domains of sport and exercise medicine/sport science, behavioural science and psychology, methodological and statistical journals, and public health or substance-use outlets. Sport science and sport medicine journals typically publish event-based or elite athlete prevalence studies, while behavioural and social science journals emphasise issues of sensitive behaviour, social desirability and response processes. Methodological journals are largely devoted to model development and validation rather than substantive prevalence estimation (albeit producing prevalence estimations as a 'by-product' of model testing), whereas public health and addiction journals feature more prominently in early work and studies of fitness or recreational sport. This dispersion reflects considerable methodological sophistication but weak disciplinary consolidation, with parallel research communities that are only partially connected. Although the populations studied have diversified over time from elite athletes to recreational, fitness, and ultra-endurance athletes, elite sport continues to function as the dominant normative reference point for interpretation and policy relevance.

Geographically, the field remains predominantly European, driven in particular by German and Dutch research groups. Contributions from outside Europe are comparatively rare and tend to be a one-off context-specific application rather than programmatic. While this reflects Europe's leading role in both methodological innovation and anti-doping policy, it also exposes a Eurocentric bias that limits the cultural and linguistic diversity of the evidence base. The way sensitive questions such as doping are framed and the extent to which respondents trust researchers or institutions, is profoundly shaped by cultural context and language [96]. Limited participation from non-European regions may therefore restrict understanding of how IEM-based instruments should be adapted to diverse populations to ensure conceptual, ethical, and linguistic equivalence.

Our bibliometric and narrative analyses identified several instances in which identical datasets and findings were disseminated across multiple publications, sometimes in different languages or formats. Although such practices may increase accessibility, they complicate evidence synthesis by increasing the likelihood of double-counting and by artificially amplifying measures of scholarly impact. Re-analyses were most commonly motivated by efforts to refine IEM-based prevalence estimation and to model alternative patterns of survey instruction noncompliance. Consequently, multiple prevalence estimates are frequently reported from the same underlying samples. A clear example is the sequence of studies by Cruyff et al. [78] and Sayed et al. [32], which progressively extended the CM to account for self-protective responding and inattentive random responding, respectively. Similarly, Ulrich et al. [18] re-analysed data originally reported by Petróczy et al. [67], producing substantially different prevalence estimates and contributing to an ongoing methodological debate regarding the interpretation of earlier findings derived from the same populations and events [18,62,67].

Although these analytical refinements are scientifically defensible, they generate multiple, sometimes divergent estimates from identical datasets, complicating public-facing communication about doping prevalence. This challenge is evident in recent scholarly exchanges concerning the interpretation and policy relevance of such estimates [97–99]. While scenario-based modelling of noncompliance enhances methodological insight and empirical testability, interpretation depends critically on understanding the underlying behavioural assumptions and model specifications. Absent this contextualisation, successive refinements may appear inconsistent or even suspect to researchers, policymakers, regulators, and media audiences. This dynamic risks undermining confidence in IEM and may erode trust among practitioners who rely on prevalence estimates for risk assessment, adjudication, and evaluation of anti-doping policy effectiveness [1].

Taken together, the temporal and bibliometric evidence points to a clear epistemic evolution. An initial phase of estimation optimism and replication gave way to growing awareness of construct overlap and potential inflation, followed more recently by a period of reflexivity marked by bias modelling, uncertainty, and reassessment of legacy estimates. Later studies increasingly foreground model assumptions, researcher degrees of freedom, and interpretive limits, explicitly challenging the treatment of prevalence figures as stable or definitive indicators. This trajectory shows both scientific maturation of the field and persistent difficulties surrounding the communication and use of IEM-based doping prevalence estimates beyond specialist audiences.

Across the corpus, *doping* functions as a metonym for multiple overlapping phenomena, ranging from elite rule violations to everyday enhancement behaviours. A number of titles signal this conceptual fluidity by referring interchangeably to doping, performance-enhancing substances, drugs or pharmacological enhancers. Only a minority explicitly specify substances (e.g., anabolic steroid), or distinguish between intentional and inadvertent use, or between physical and cognitive enhancement. Such ambiguity in the definition of doping is characteristic of the field and has been highlighted as a hindering factor in doping behaviour research [100,101] and communication [5].

5. Discussion

The combined bibliometric and narrative analyses depict a field that is methodologically innovative yet structurally constrained and epistemically fragmented. Distinct biomedical, behavioural, and sociological framings of doping correspond to separate intellectual communities, each characterised by its own methodological preferences, publication venues, and linguistic conventions. These *invisible colleges* shape how doping is studied and communicated, reinforcing parallel rather than integrated lines of inquiry.

Patterns of authorship concentration and clustering reflect both the strengths and limitations of a specialised community operating at the intersection of behavioural science, statistics, and sport ethics. Given the technical demands of IEM, such group-specific alignments are not unexpected and resemble developmental trajectories observed in other behavioural domains. For example, in the evolution of the Implicit Association Test [102–105] and in orthorexia research, where early fragmentation prompted later conceptual consolidation [106,107]. Akin to these examples, the diversity of conceptual and linguistic framings contributes to ongoing ambiguity in how *doping* is defined, operationalised, and interpreted. Variation in terminology complicates evidence synthesis and cross-study comparison and reflects broader fragmentation across publication outlets [5,101]. Citation patterns further reinforce this dynamic: studies reporting unusually high prevalence estimates tend to attract disproportionate attention, amplifying methodological debates while sometimes sidelining nuance.

A notable gap across the corpus is the absence of qualitative or mixed-method work exploring how respondents understand and engage with IEM surveys. Existing evidence suggests that comprehension, trust, and emotional responses influence data quality and may contribute to noncompliance [16,28,108,109]. Without insight into these processes, refinements in statistical modelling risk outpacing understanding of respondent behaviour. Incorporating qualitative methods such as cognitive interviews or think-aloud protocols could help distinguish between true

concealment and methodological artefacts and provide a behavioural foundation for future model development. These limitations also intersect with the field's Eurocentric orientation. Most IEM-based doping prevalence studies have been designed and interpreted within Western European contexts, raising questions about cultural transferability. Assumptions about privacy, probabilistic reasoning, and institutional trust may not hold globally, making cross-cultural and qualitative validation essential for ensuring conceptual, ethical, and measurement robustness.

5.1. Authorship Structure and Implications

Authorship analysis revealed a structurally narrow research community, with only a small group of scholars possessing expertise in both IEM methodology and doping research. Within this small community, the authorship and citation networks show a small number of densely interconnected clusters resembling "invisible colleges" [89], each aligned with particular IEM variants or analytical traditions. These communities facilitate cumulative methodological development but also risk reinforcing established paradigms and limiting cross-fertilisation. Limited dialogue between clusters may entrench methodological divides, slowing conceptual innovation. As Zuccala [110] argues, such invisible colleges persist through the practices of information users (in this case, researchers, policymakers, and critics) whose engagement patterns shape visibility, influence, and impact across the field.

This limited pool of experts also raises challenges for expert peer review. Repeated reliance on the same experts risks intellectual insularity, while broadening the reviewer base often brings in specialists who understand either the modelling or the doping context, but not both. These constraints reduce the depth of methodological and contextual evaluation and underscore the need for interdisciplinary collaboration, methodological cross-training, and greater transparency in reviewer expertise.

Gender composition adds another layer to the field's structural dynamics. Whereas women have been comparatively well represented in the broader anti-doping research landscape [88,111,112], IEM-based prevalence research remains predominantly male. This likely reflects the disciplinary origins of IEM work within quantitative and mathematical traditions, which remain male-dominated globally. Such imbalances may subtly influence the types of questions pursued and the epistemic styles privileged, reinforcing methodological orientations that favour formal modelling and quantification over more contextual or relational approaches to understanding doping behaviour.

Collectively, these structural features point to a research landscape that is productive but fragile: reliant on a small and interconnected community, shaped by disciplinary and gendered trajectories, and susceptible to epistemic insularity. Strengthening interdisciplinary collaboration, broadening methodological repertoires, and promoting more inclusive and transparent review practices will be essential for sustaining rigour and innovation in IEM-based doping prevalence research.

5.2. The Interpretive Scope and Boundaries of 'Evidence'

The synthesis presented here allows for qualified statements about the relative strength and convergence of IEM-based doping prevalence estimates across models and analytical approaches. By mapping where multiple primary studies align, and where evidence remains sparse or reliant on reanalysis, the review identifies prevalence ranges that are more, or less, strongly supported within the existing literature. In this sense, the analysis clarifies patterns of evidentiary robustness rather than producing a single summary estimate. At the same time, the synthesis does not permit claims about subgroup-specific prevalence, differences by athlete level, or the identification of a definitive or "true" rate of doping. Nor does it adjudicate between competing definitions of doping used across studies.

Citation analysis revealed a consistent asymmetry favouring studies with higher or more dramatic prevalence estimates, which tend to attract disproportionate academic and media attention. This pattern suggests that visibility and influence within the field may be shaped as much by the perceived newsworthiness of findings as by methodological innovation or quality. While such

attention can raise awareness of doping as a social issue, it risks overshadowing more nuanced or conservative studies that may offer greater validity.

Moreover, the probabilistic nature of IEM-derived estimates makes them vulnerable to misinterpretation by audiences unfamiliar with indirect estimation principles. Without appropriate context, these figures may be misconstrued as direct evidence of doping rates rather than statistical inferences. Authors, reviewers, and editors therefore share responsibility for clear and transparent communication—providing interpretive guidance, confidence intervals, and explicit caveats to prevent sensationalism and misuse of complex quantitative data.

5.3. The Impact of Duplicate Publications and Re-Analyses

Duplicate publications and secondary data re-analyses pose distinct challenges for evidence synthesis and bibliometric evaluation. Duplicate publications, often justified on grounds of audience reach or language accessibility, complicate systematic reviews by increasing the risk of data duplication and distort bibliometric indicators by inflating publication and citation counts. Secondary (re-)analyses, while scientifically valuable, introduce additional interpretive complexity when multiple, equally plausible prevalence estimates are derived from the same dataset. In IEM-based doping prevalence research, such re-analyses typically arise from refined assumptions regarding survey instruction noncompliance or response validity. These methodological iterations contribute to improved model robustness and theoretical clarity, but they may blur the boundary between legitimate refinement and secondary data analysis [113] and questionable academic practices such as selective reporting, p-hacking, HARKing, RHARKing, CHARKing or salami slicing [114,115]. Clearer reporting standards for explicitly documenting data provenance, analytical rationale, pre-registration and adherence to open-science principles [116,117] are therefore necessary to distinguish genuine methodological advancement from ethically problematic redundancy.

5.4. Practical Implications

These findings have several implications for researchers, practitioners, and journal editors operating at the intersection of sport science, behavioural research, and anti-doping policy. For researchers, the results underscore the importance of transparent reporting and reflexivity when applying IEM. Clear documentation of data provenance, analytical assumptions, and the rationale for re-analysis should become standard practice to reduce duplication bias and enhance interpretive clarity. Cross-disciplinary training integrating behavioural science, psychometrics, and sport ethics may further strengthen methodological and contextual competence. For practitioners and anti-doping organisations, a nuanced understanding of IEM-derived prevalence estimates is essential. As these estimates are probabilistic rather than diagnostic, they are best suited to informing population-level strategy rather than individual-level judgement. Training and communication materials should therefore emphasise the interpretive limits of IEM outputs and situate them within broader evidence frameworks, including testing statistics, education programme indicators, and sociocultural data.

For journal editors and reviewers, diversifying the peer-review process is critical. Engaging reviewers with complementary methodological and applied expertise (rather than relying on a narrow group of specialists) can mitigate intellectual clustering and promote balanced evaluation. Editors may also consider requiring explicit statements on data re-use, analytical transparency, and open-science compliance. Across all stakeholder groups, greater attention to cultural and linguistic sensitivity in the design, analysis, and dissemination of IEM-based research is essential to enhance trust, data quality, and the ethical integrity of future doping prevalence studies.

5.5. Study Limitations and Future Directions

This study has limitations that need to be acknowledged. Bibliometric analyses were possible only for outputs indexed in Web of Science and Scopus, introducing potential database-selection bias toward English-language, higher-impact journals and excluding studies found in regional or

non-indexed outlets. Likewise, the narrative component focused on English-language publications, limiting interpretive depth for studies available only in other languages. In evaluating evidentiary strength, estimates were not disaggregated by athlete sport involvement level or by definition of doping but rather, it focused on IEM models and whether the analysis assumed and accounted for noncompliance in some form of secondary analysis. Readers needing detailed quantitative synthesis should consult Sagoe et al. [26].

The authors of this review contributed to some of the analysed outputs, which presents challenges to complete detachment. This was mitigated through objective bibliometric procedures, transparent inclusion criteria, and involvement of bibliometric experts without prior publications in doping. Nonetheless, our narrative interpretations inevitably reflect our own epistemic orientations. We therefore foreground positionality as part of reflexive and mixed-method scholarship. The authorship team's disciplinary balance, gender diversity, and varied involvement in anti-doping contributes to epistemic breadth rather than bias.

Future work should expand bibliometric coverage beyond Web of Science and Scopus to include regional and language-specific databases, reducing indexation bias and offering a more global picture of IEM use. Multilingual narrative analyses would further illuminate how conceptualisations and reporting practices vary across cultural contexts. We feel that this is an important consideration given that the framing of sensitive behaviours like doping is shaped by cultural norms and moral discourses.

Greater attention to participant experience with IEM is also needed. Although IEM are designed to protect anonymity, their validity depends on respondents' comprehension, trust, and motivation. Evidence from sensitive survey research highlights the role of misunderstanding, self-protection, or disengagement in driving noncompliance. Cognitive interviewing, think aloud protocols, and cross-cultural piloting would clarify these behavioural processes and improve model robustness. Understanding and modelling instruction noncompliance remains a priority, given its central role in biasing prevalence estimates. Future studies should triangulate theoretical models of noncompliance with empirical behavioural data—via experiments, response time analysis, or behavioural tracking—to refine correction procedures and strengthen interpretive validity.

Cross-community collaboration remains essential for mitigating the intellectual insularity sustained by methodological and disciplinary "invisible colleges". Interdisciplinary research bridging statistical modelling with sport science, behavioural science, and ethics would broaden interpretive perspectives. Joint authorship, shared data repositories, open methodological documentation, and interdisciplinary symposia could promote methodological learning and reduce fragmentation. The field would also benefit from specialised methodological guidelines for evidence synthesis in IEM-based doping research. Current systematic-review frameworks are not well equipped to handle duplicated outputs, re-analyses, and cross-model heterogeneity. Developing consensus-based standards analogous to PRISMA extensions tailored to complex or IEM would enhance transparency, reduce duplication bias, and improve comparability across studies.

6. Conclusions

This study complements existing doping prevalence estimates and their systematic and meta-analytic synthesis [26] by situating IEM within their intellectual, social, and methodological ecosystems. Understanding these dynamics helps placing IEM-based prevalence estimates into policy, practical and research context, and cautions against the 'higher must be more truthful' heuristic, selective citation, and overinterpretation. Over time, IEM-based doping prevalence research has evolved from early prevalence reporting toward greater methodological reflexivity and specialisation. Despite increased internal coherence and visibility, the field remains constrained by Eurocentrism, intellectual clustering, and uneven interpretive standards.

To date, IEM-based doping prevalence research has focused almost exclusively on either statistical method development or straightforward application to generate prevalence estimates, with a notable gap in studies examining how athletes perceive, experience, and respond to sensitive

doping questions within IEM survey environments. Attention to this gap is critical for understanding survey instruction noncompliance and for informing robust post-data-collection adjustment strategies. Future priorities include cultural adaptation of IEM instruments, qualitative investigation of behavioural dynamics, triangulation with empirical data, and the development of method-specific standards for synthesising IEM-derived prevalence estimates, including adjustments for survey instruction noncompliance. Addressing these limitations is essential to ensure that prevalence estimates are not only statistically robust, but ethically sound, culturally grounded, and fit for informing anti-doping policy and governance. Given the methodological complexity of IEM, new users are strongly encouraged to collaborate with experienced experts. Policymakers and practitioners should also attend closely to the intended purpose of prevalence studies, distinguish method development from prevalence estimation, and exercise particular caution when interpreting estimates derived from studies focusing on methodological improvement and validation.

Supplementary Materials: The following supporting information can be downloaded at: Preprints.org, Table S1: Distribution of the included scientific journal articles by journals.; Table S2: Distribution of the outputs indexed in WoS by research topics.; Table S3: Scientific Impact assessment.; Table S4: Evidentiary summary table with references.; Table S5: List of authors involved in IEM application to doping prevalence estimation.

Author Contributions: Conceptualization: A.P., D.S., A.K. and S.S.; methodology: A.P., D.S., A.K., S.S.; formal analysis: A.P., A.K., S.S.; investigation: D.S., R.C., A.V., M.C., P.vdH.; writing—original draft preparation: A.P., D.S., A.K., S.S.; writing—review and editing: A.P., D.S., O.dH., R.C., M.C., P.vdH.; visualization: A.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: A.P., D.S., M.C., and O.H. have been members of the Prevalence Working Group (PWG) of the World Anti-Doping Agency (WADA) since 2021. PWG members receive no salary for their work for WADA but are entitled to expenses covered, and receive honorarium for formal meetings of up to five days per year for preparation. Other authors report no competing interests.

Abbreviations

The following abbreviations are used in this manuscript:

CHARKing	Cherry-picking significant results.
CW	Crosswise model
FR	Forced Response model
FWCI	Field Weighted Citation Index
IEM	Indirect estimation models
RHARKing	Retrieving hypotheses from post hoc literature searches.
SHARKing:	Suppressing unsupported a priori hypotheses
SSC	Single Sample Count model
UQM	Unrelated Question model

References

1. Petróczy, A. Numbers do not lie, but they can mislead: rethinking what doping prevalence statistics really mean. *J Sports Med. Phys. Fitness* **2025**, *65*, 835–838. <https://doi.org/10.23736/S0022-4707.25.17281-2>
2. Gleaves, J.; Petróczy, A.; Folkerts, D.; De Hon, O.; Macedo, E.; Saugy, M.; Cruyff, M. Doping prevalence in competitive sport: evidence synthesis with “best practice” recommendations and reporting guidelines from

- the WADA Working Group on Doping Prevalence. *Sports Med.* **2021**, *51*, 1909–1934. <https://doi.org/10.1007/s40279-021-01477-y>
3. Lockett, I.; Blank, C.; Patterson, L.; Westmattmann, D.; Lux, D.; Petróczy, A. From violation to stigma: a literature review of athletes' lived experiences following anti-doping sanctions. *Front. Sports Act. Living* **2026**, *8*, 1651135. <https://doi.org/10.3389/fspor.2026.1651135>
 4. Lockett, I.; Exner, J.; Pummell, E.; Petróczy, A. Mapping doping-related criminal legislation together: An informed stakeholder consultation. *Perform. Enhanc. Health* **2026**, *14*, 100413. <https://doi.org/10.1016/j.peh.2026.100413>
 5. Grimes, H.; Cox, L. T. J. Talking dirty: Anti-doping's stigmatizing rhetoric and its impact on the unintentional doper. *Perform. Enhanc. Health* **2026**, *14*, 100412. <https://doi.org/10.1016/j.peh.2025.100412>
 6. Arnab, R. Randomized Response Techniques: Early developments. In *Indirect methods of data collection and analysis from surveys*. Springer Nature Singapore, Singapore **2025**, pp. 1-30.
 7. Le, T.N.; Lee, S.M.; Tran, P.L.; Li, C.S. Randomized response techniques: a systematic review from the pioneering work of Warner (1965) to the present. *Mathematics* **2023**, *11*, 1718. <https://doi.org/10.3390/math11071718>
 8. Lensvelt-Mulders, G.J.; Hox, J.J.; Van der Heijden, P.G.; Maas, C.J. Meta-analysis of randomized response research: thirty-five years of validation. *Sociol. Methods Res.* **2005**, *33*, 319–348.
 9. Boruch, R.F. Assuring confidentiality of responses in social research: a note on strategies. *Am. Sociol.* **1971**, *6*, 308–311.
 10. Kuk AY. Asking sensitive questions indirectly. *Biometrika* **1990**, *77*, 436–438.
 11. Greenberg, B.G.; Abul-Ela, A.L.; Simmons, W.R.; Horvitz, D.G. The unrelated question randomized response model: theoretical framework. *J. Am. Stat. Assoc.* **1969**, *64*, 520-539. <https://doi.org/10.1080/01621459.1969.10500991>
 12. Horvitz, D.G.; Shah, B.V.; Simmons, W.R. The unrelated question randomized response model. *Social Stat. Sect. Proc. Am. Stat. Assoc.* **1967**, 65–72.
 13. Yu, J.W.; Tian, G-L.; Tang, M-L. Two new models for survey sampling with sensitive characteristic: design and analysis. *Metrika* **2008**, *67*, 251–263. <https://doi.org/10.1007/s00184-007-0131-x>
 14. Petróczy, A.; Nepusz, T.; Cross, P.; Taft, H.; Shah, S.; Deshmukh, N.; Schaffer, J., Shane, M., Adesanwo, C., Barker, J., Naughton, D.P. New non-randomised model to assess the prevalence of discriminating behaviour: a pilot study on mephedrone. *Subst. Abuse Treat. Prev. Policy* **2011**, *6*, 20. <https://doi.org/10.1186/1747-597X-6-20>
 15. Nepusz, T.; Petróczy, A.; Naughton, D.P.; Epton, T.; Norman, P. Estimating the prevalence of socially sensitive behaviors: attributing guilty and innocent noncompliance with the single sample count method. *Psychol Methods* **2014**, *19*, 334–355. <https://doi.org/10.1037/a0034961>
 16. Boeije, H.; Lensvelt-Mulders, G. Honest by chance: a qualitative interview study to clarify respondents' (non-) compliance with computer-assisted randomized response. *Bull. Sociol. Methodol.* **2002**, *75*, 24–39. <https://doi.org/10.1177/075910630207500104>
 17. Pielke, R. Assessing doping prevalence is possible. So what are we waiting for?. *Sports Med.* **2018**, *48*, 207–209. <https://doi.org/10.1007/s40279-017-0792-1>
 18. Ulrich, R.; Cléret, L.; Comstock, R.D.; Kanayama, G.; Simon, P.; Pope, H.G. Jr. Assessing the prevalence of doping among elite athletes: An analysis of results generated by the Single Sample Count method versus the Unrelated Question Method. *Sports Med. Open* **2023**, *9*(1), 112. <https://doi.org/10.1186/s40798-023-00658-5>
 19. Ibbett, H.; Dorward, L.J.; Kohi, E.M.; Jones, J.P.; Sankeni, S.; Kaduma, J.; Mchomvu, J.; Mawenya, R.; St. John, F.A. Topic sensitivity still affects honest responding, even when specialized questioning techniques are used. *Conserv. Sci. Pract.* **2023**, *5*, e12927. <https://doi.org/10.1111/csp2.12927>
 20. Clark, S.J.; Desharnais, R.A. Honest answers to embarrassing questions: detecting cheating in the randomized response model. *Psychol. Methods* **1998**, *3*, 160–168. <https://doi.org/10.1037/1082-989X.3.2.160>
 21. Ostapczuk, M.; Much, J.; Moshagen, M. Improving self-report measures of medication non-adherence using a cheating detection extension of the randomised-response-technique. *Stat. Methods Med. Res.* **2011**, *20*, 489–503. <https://doi.org/10.1177/0962280210372843>

22. Liu, Y.; Tian, G.L. A variant of the parallel model for sample surveys with sensitive characteristics. *Comput. Stat. Data Anal.* **2013**, *67*, 115–135. <https://doi.org/10.1016/j.csda.2013.05.003>
23. Tian, GL. A new non-randomized response model: the parallel model. *Stat. Neerl.* **2014**, *68*, 293–323. <https://doi.org/10.1111/stan.12034>
24. Heck, D.W.; Hoffmann, A.; Moshagen, M. Detecting nonadherence without loss in efficiency: a simple extension of the crosswise model. *Behav Res Methods* **2018**, *50*, 1895–1905. <https://doi.org/10.3758/s13428-017-0957-8>
25. Reiber, F.; Pope, H.; Ulrich, R. Cheater detection using the unrelated question model. *Sociol. Methods Res.* **2023**, *52*, 389–411. <https://doi.org/10.1177/0049124120914919>
26. Sagoe, D.; Cruyff, M.; Chegeni, R.; Veltmaat, A.; Kiss, A.; Soós, S.; De Hon, O., Van der Heijden, P., Petróczi, A. Exploring doping prevalence in sport from indirect estimation models: a systematic review and meta-bibliometric analysis. preprint, **2024** <https://doi.org/10.21203/rs.3.rs-4104397/v1>
27. De Schrijver, A. Sample survey on sensitive topics: Investigating respondents' understanding and trust in alternative versions of the randomized response technique. *J Res. Pract.* **2012**, *8*(1), 1–17. <http://jrp.icaap.org/index.php/jrp/article/view/277/250>
28. Jerke, J.; Johann, D.; Rauhut, H.; Thomas K. Too sophisticated even for highly educated survey respondents? A qualitative assessment of indirect question formats for sensitive questions. *Surv Res Methods* **2019**, *13*, 319–351. <https://doi.org/10.18148/srm/2019.v13i3.7453>
29. Walzenbach, S.; Hinz, T. Puzzling answers to crosswise questions: Examining overall prevalence rates, response order effects, and learning effects. *Surv. Res. Methods* **2023**, *17*(1), 1–13. <https://doi.org/10.18148/srm/2023.v17i1.8010>
30. Creswell, J.W., Clark, V.L.P. *Designing and conducting mixed methods research*. Sage, London, **2017**.
31. Pitsch, W.; Christiansen, A. V. Dope stereotypes: When perception runs south and prevalence points north. *Perform. Enhanc. Health* **2026**, *14*(1), 100398. <https://doi.org/10.1016/j.peh.2025.100398>
32. Sayed, K. H.; Cruyff, M. J.; Petróczi, A.; Van der Heijden, P. G. The Extended Crosswise Model adjusted for random answering. *J. Surv. Stat. Methodol.* **2026** (in press), Available since 2024 at arXiv preprint arXiv:2412.09506.
33. Schu, K.; Haller, N. Cheating and doping in chess – A survey among 1,924 German club players using the Randomized Response Technique. *Perform. Enhanc. Health* **2025**, *13*, 100344. <https://doi.org/10.1016/j.peh.2025.100344>
34. Musch, J.; Plessner, H. A randomized response investigation of the prevalence of doping. **2002**. Unpublished manuscript used with authors' permission.
35. Plessner, H.; Musch, J. Wie verbreitet ist Doping im Leistungssport? Eine www Umfrage mit Hilfe der Randomized-Response-Technik [How widespread is doping in competitive sports? A www survey using the randomized response technique]. In: Strauß B, editor. *Expertise im sport*. BPS, Cologne **2002**. p. 78–9.
36. Pitsch, W., Emrich, E., Klein, M. Zur Häufigkeit des Dopings im Leistungssport: Ergebnisse eines www-surveys [On the frequency of doping in high-performance sport: results of a www survey]. *Leipziger Sportwissenschaftliche Beiträge* **2005**, *46*, 63–77.
37. Simon, P.; Striegel, H.; Aust, F.; Dietz, K.; Ulrich, R. Doping in fitness sports: estimated number of unreported cases and individual probability of doping. *Addiction* **2006**, *101*(11), 1640–1644. <https://doi.org/10.1111/j.1360-0443.2006.01568.x>
38. Pitsch, W.; Emrich, E.; Klein M. Doping in elite sports in Germany: results of a www survey. *Eur. J Sport Soc.* **2007**, *4*, 89–102. <https://doi.org/10.1080/16138171.2007.11687797>
39. Pitsch, W.; Maats, P.; Emrich, E. Zur Häufigkeit des Dopings im deutschen Spitzensport [On the frequency of doping in German elite sport]. *Magazin Forschung* **2009**, 15–19.
40. Pitsch, W.; Maats, P.; Emrich, E. Zur Häufigkeit des Dopings im deutschen Spitzensport–eine Replikationsstudie [On the frequency of doping in German elite sport–a replication study]. In: Emrich E, Pitsch W, editors. *Sport und Doping: zur Analyse einer antagonistischen Symbiose*. Peter Lang, Frankfurt **2009**. pp. 19–36.
41. Pitsch, W.; Maats, P.; Emrich, E. On the frequency of doping in top German sport - a replication study. E. Emrich, W Pitsch (Eds) *Sport and doping. For the analysis of an antagonistic symbiosis*, **2009**, pp 19–36.

42. 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), 230-232. <https://doi.org/10.1016/j.drugalcdep.2009.07.026>
43. Stamm, H.; Stahlberger, M.; Gebert, A.; Lamprecht, M.; Kamber, M.; Schweiz, A. Supplemente, Medikamente und Doping im Freizeitsport. *Schweizerische Zeitschrift für Sportmedizin und Sporttraumatologie* **2011**, *59*(3), 122.
44. Pitsch, W.; Emrich, E. The frequency of doping in elite sport: results of a replication study. *Int. Rev. Sociol. Sport* **2012**, *47*, 559–580. <https://doi.org/10.1177/1012690211413969>
45. Striegel, H. Doping im Breiten- und Freizeitsport. In: Vieweg, K. (ed.). *Akzente des Sportrechts* (1st ed.): Duncker & Humblot, Berlin, **2012**, pp. 31-42.
46. Breuer, C.; Hallmann, K. Dysfunktionen des Spitzensports: doping, match-fixing und gesundheitsgefährdungen aus sicht von bevölkerung und athleten. Bundesinst. für Sportwissenschaft. **2013**, <https://fis.dshs-koeln.de/en/publications/dysfunktionen-des-spitzensports-doping-match-fixing-und-gesundheit>
47. Dietz, P.; Ulrich, R.; Dalaker, R.; Striegel, H.; Franke, A. G.; Lieb, K.; Simon, P. Associations between physical and cognitive doping – a cross-sectional study in 2997 triathletes. *PLoS One* **2013**, *8*, 11. <https://doi.org/10.1371/journal.pone.0078702>
48. James, R. A.; Nepusz, T.; Naughton, D. P.; 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. <https://doi.org/10.1016/j.psychsport.2012.08.003>
49. Nakhaee, M. R.; Pakravan, F.; Nakhaee, N. Prevalence of use of anabolic steroids by bodybuilders using three methods in a city of Iran. *Addict. Health* **2013**, *5*(3-4), 77.
50. Pitsch, W.; Emrich, E.; Frenger, M. Doping im Breiten- und Freizeitsport. Zur Überprüfung von Hypothesen mittels RRT-gewonnener Daten. H Kempf, S Nagel, H Dietl (Eds) *Im Schatten der Sportwirtschaft*. Hofmann. **2013**.
51. Anti-Doping Agency of Serbia. Who is your team? The Importance of "Sport Entourage" for Sport Fellows of Serbia - Recommendations to Ministry of Youth and Sports. Belgrade: Anti-Doping Agency of Serbia, **2014**
52. Stubbe, J. H.; Chorus, A. M.; Frank, L. E.; De Hon, O., Van der Heijden, P. G. Prevalence of use of performance enhancing drugs by fitness centre members. *Drug Testing Anal.* **2014**, *6*(5), 434-438. <https://doi.org/10.1002/dta.1525>
53. Duiven, E.; De Hon, O. De Nederlandse topsporter en het anti-dopingbeleid 2014 - 2015 [The Dutch elite athlete and anti-doping policy 2014 - 2015. Capelle aan den IJssel: Anti-Doping Authority Netherlands **2015**. [International summary retrieved from: [retrieved from. https://www.dopingautoriteit.nl/media/files/2015/The_Dutch_elite_athlete_and_the_anti-doping_policy_2014-2015_international_summary_DEF.pdf](https://www.dopingautoriteit.nl/media/files/2015/The_Dutch_elite_athlete_and_the_anti-doping_policy_2014-2015_international_summary_DEF.pdf)]
54. Backhouse, S.; Whitaker, L.; McKenna, J.; Beggs, C.; Watkins, S.; Nunn, R.; Petroczi, A. Schoolboy supplement use behaviours and doping vulnerability. **2016** <https://eprints.leedsbeckett.ac.uk/id/eprint/7554/1/SchoolboySupplementUseBehavioursAndDopingVulnerabilityPV-BACKHOUSE.pdf>
55. 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), 1965-1969. <https://doi.org/10.1080/02640414.2016.1149214>
56. Frenger, M.; Pitsch, W.; Emrich, E. Sport-induced substance use—An empirical study to the extent within a German Sports Association. *PloS One* **2016**, *11*, 10. <https://doi.org/10.1371/journal.pone.0165103>
57. Schröter, H.; Studzinski, B.; Dietz, P.; Ulrich, R.; Striegel, H.; Simon, P. A Comparison of the cheater detection and the unrelated question models: a randomized response survey on physical and cognitive doping in recreational triathletes. *PloS One* **2016**, *11*, 5. <https://doi.org/10.1371/journal.pone.0155765>
58. Fincoeur, B.; Pitsch, W. Omgaan met sociale wenselijkheid: Inschatting van de dopingprevalentie aan de hand van de Randomized Response Technique. *Panopticon J. Crim. Law Criminol. Crim. Justice* **2017**, *38*(5), 376-386. <https://iris.unil.ch/handle/iris/57463>

59. Franke, A. G.; Dietz, P.; Ranft, K.; Balló, H.; Simon, P.; Lieb, K. The use of pharmacologic cognitive enhancers in competitive chess. *Epidemiol.* **2017**, *28*(6), e57-e58. <http://doi.org/10.1097/EDE.0000000000000737>
60. Elbe, A. M.; Pitsch, W. Doping prevalence among Danish elite athletes. *Perform. Enhanc. Health* **2018**, *6*(1), 28-32. <https://doi.org/10.1016/j.peh.2018.01.001>
61. Pitsch, W. Assessing and explaining the doping prevalence in cycling. In B. Fincoeur, J. Gleaves, F. Ohl (Eds.) *Doping in cycling: Interdisciplinary perspectives*. Routledge, Abingdon. **2018**, pp 13-30.
62. Ulrich, R.; Pope, H. G.; Cléret, L.; Petróczi, A.; Nepusz, T.; Schaffer, J.; Kanayama, G.; Comstock, R. D.; Simon, P. Doping in two elite athletics competitions assessed by randomized-response surveys. *Sports Med.* **2018**, *48*(1), 211-219. <https://doi.org/10.1007/s40279-017-0765-4>
63. Boardley, I. D.; Smith, A. L.; Ntoumanis, N.; Gucciardi, D. F.; Harris, T. S. Perceptions of coach doping confrontation efficacy and athlete susceptibility to intentional and inadvertent doping. *Scand. J. Med. Sci. Sports* **2019**, *29*(10), 1647-1654. <https://doi.org/10.1111/sms.13489>
64. Seifarth, S.; Dietz, P.; Disch, A. C.; 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* **2019**, *7*(12), 241. <https://doi.org/10.3390/sports7120241>
65. Heller, S.; Ulrich, R.; Simon, P.; Dietz, P. (2020). Refined analysis of a cross-sectional doping survey among recreational triathletes: Support for the nutritional supplement gateway hypothesis. *Front. Psychol.* **2020**, *11*, 561013. <https://doi.org/10.3389/fpsyg.2020.561013>
66. Nilaweera, A.; Nadishani, U.; Nipunya, G.; Wijekoon, N. 369 Knowledge, attitude and usage of doping drugs among national level athletes in Sri Lanka. *Br. J Sports Med.* **2020**, *54*(Suppl 1), A150-A150. <http://dx.doi.org/10.1136/bjsports-2021-IOC.327>
67. Balk, L.; Dopheide, M. Dopinggebruik in de Nederlandse topsport [Doping use in Dutch elite sport]. Utrecht: Mulier Institute; **2021**. Available at: <https://www.mulierinstituut.nl/publicaties/25952/doping-in-dutch-elite-sports/>
68. Petróczi, A.; Cruyff, M.; De Hon, O.; Sagoe, D.; Saugy, M.O. Hidden figures: revisiting doping prevalence estimates reported for two major international sport events in Ulrich et al. (2018) in the context of further empirical evidence and the extant literature. *Front. Sports Act. Living* **2022**, *4*, 1017329. <https://doi.org/10.3389/fspor.2022.1017329>
69. Hilkens, L.; Cruyff, M.; Woertman, L.; Benjamins, J.; Evers, C. (2021). Social media, body image and resistance training: Creating the perfect 'Me' with dietary supplements, anabolic steroids and SARM's. *Sports Med. Open* **2021**, *7*(1), 1-13. <https://doi.org/10.1186/s40798-021-00371-1>
70. Heyes, A. R. Psychosocial factors facilitating use of performance and cognitive enhancing drugs in sport and education (Doctoral dissertation, University of Birmingham). **2022**.
71. Pitsch, W. (2022). Doping in recreational sport as a risk management strategy. *J. Risk Financ. Manag.* **2022**, *15*(12), 574. <https://doi.org/10.3390/jrfm15120574>
72. Reiber, F.; Schnuerch, M.; Ulrich, R. Improving the efficiency of surveys with randomized response models: A sequential approach based on curtailed sampling. *Psychol. Methods* **2022**, *27*(2), 198-211. <https://doi.org/10.1037/met0000353>
73. Sayed, K. H.; Cruyff, M. J.; Van der Heijden, P. G.; Petróczi, A. Refinement of the extended crosswise model with a number sequence randomizer: Evidence from three different studies in the UK. *Plos One* **2022**, *17*(12), e0279741. <https://doi.org/10.1371/journal.pone.0279741>
74. World Anti-Doping Agency. Doping Prevalence Working Group (Petróczi A, De Hon O, Saugy M, Cruyff M, Sagoe D, Gleaves J) interim report (Unpublished report). Montreal: Canada; **2022**.
75. Abdulrazzaq, Z.; Tareq, A. The Psychosomatic Reflection of AAS (Androgenic Anabolic Steroid) Usage between Bodybuilders in Baghdad Gyms. *J. ReAtt. Ther. Dev. Divers.* **2023**, *6*(2s), 224-232. <https://www.jrtdd.com/index.php/journal/article/view/287>
76. Balk, L.; Dopheide, M.; Cruyff, M.; Erik, D.; De Hon, O. Doping prevalence and attitudes towards doping in Dutch elite sports. *Sci. J. Sport Perform.* **2023**, *2*(2), 132-143. <https://doi.org/10.55860/BCUQ4622>

77. Christiansen, A. V.; Frenger, M.; Chirico, A.; Pitsch, W. Recreational athletes' use of performance-enhancing substances: Results from the first European Randomized Response Technique Survey. *Sports Med. – Open* **2023**, *9*(1), 1-17. <https://doi.org/10.1186/s40798-022-00548-2>
78. Cruyff, M. J.; Sayed, K. H.; Petróczi, A.; Van der Heijden, P. G. The one-sayers model for the Extended Crosswise design. *J. R. Stat. Soc. A* **2024**, *187*(4), 882–899. <https://doi.org/10.1093/jrssa/qnae009>
79. Robach, P.; Trebes, G.; Buisson, C.; Mechin, N.; Mazzarino, M.; Garribba, F.; Roustit, M.; Quesada, J.L.; Lefèvre, B.; Giardini, G.; De Seigneux, S.; Botre, F.; Bouzat, P. Prevalence of drug use in ultra-endurance athletes. *Med. Sci. Sports Exerc.* **2024**, *56*(5), 828-838. <https://doi.org/10.1249/mss.0000000000003374>
80. Sayed, K. H.; Cruyff, M. J.; Van der Heijden, P. G. The analysis of randomized response “ever” and “last year” questions: A non-saturated Multinomial model. *Behav. Res. Methods* **2024**, *56*(3), 1335-1348. <https://doi.org/10.3758/s13428-023-02096-3>
81. Sayed, K. H.; Cruyff, M. J.; Van Der Heijden, P. G. Modeling evasive response bias in Randomized Response: Cheater detection versus self-protective no-saying. *Psychometrika* **2024**, *89*(4), 1261-1279. <https://doi.org/10.1007/s11336-024-10000-x>
82. Ding, Y.; Zhang, G.; Chambers, T.; Song, M.; Wang, X.; Zhai, C. Content-based citation analysis: the next generation of citation analysis. *J Assoc Inf Sci Technol* **2014**, *65*, 1820–1833. <https://doi.org/10.1002/asi.23256>
83. Garfield, E. Can citation indexing be automated? In Stevens, ME, Giuliano VE, Heilprin LB, editors. Statistical association methods for mechanized documentation. Symposium proceedings. Washington: National Bureau of Standards; **1964**. pp. 189–192.
84. Peroni, S.; Shotton, D. FaBiO and CiTO ontologies for describing bibliographic resources and citations. *J Web Semant.* **2012**, *17*, 33–43. <https://doi.org/10.1016/j.websem.2012.08.001>
85. Csardi, G., Nepusz, T. The igraph software package for complex network research. *Int J Complex Syst* **2006**, *1695*, 1–9.
86. Blondel, V.D.; Guillaume, J.L.; Lambiotte, R.; Lefebvre E. Fast unfolding of communities in large networks. *J Stat Mech Theory Exp* **2008**. P10008. <https://doi.org/10.1088/1742-5468/2008/10/P10008>
87. Palmateer, N.; Kimber, J.; Hickman, M.; Hutchinson, S.; Rhodes, T.; Goldberg, D. (2010). Evidence for the effectiveness of sterile injecting equipment provision in preventing hepatitis C and human immunodeficiency virus transmission among injecting drug users: A review of reviews. *Addiction* **2010**, *105*(5), 844–859. <https://doi.org/10.1111/j.1360-0443.2009.02888.x>
88. Kiss, A., Soós, S., Petróczi, A. Impact as equalizer: the demise of gender-related differences in anti-doping research. *Scientometrics* **2024**, *129*, 4071–4108. <https://doi.org/10.1007/s11192-024-05094-0>
89. Crane, D. *Invisible colleges: Diffusion of knowledge in scientific communities*. University of Chicago Press. **1972**.
90. De Solla Price, D. J. *Little science, big science*. Columbia University Press. **1963**.
91. Granovetter, M. S. The strength of weak ties. *Am. J Sociol.* **1973**, *78*(6), 1360–1380. <https://doi.org/10.1086/225469>
92. Burt, R. S. *Structural holes: The social structure of competition*. Harvard University Press. **1992**.
93. Burt, R. S. Structural holes and good ideas. *Am. J Sociol.* **2004**, *110*(2), 349–399. <https://doi.org/10.1086/421787>
94. Kretschmer, H. Coauthorship networks of invisible colleges and institutionalized communities. *Scientometrics* **1994**, *30*(1), 363–369. <https://doi.org/10.1007/BF02017234>
95. Nayak, T.K. A review of rigorous randomized response methods for protecting respondent's privacy and data confidentiality. Washington, D.C.: U.S. Census Bureau; **2021**. Available at: <https://www.census.gov/content/dam/Census/library/working-papers/2020/adrm/RRS2020-06.pdf>
96. Sagoe, D.; Cruyff, M.; Spendiff, O.; Chegeni, R.; De Hon, O.; Saugy, M.; Van der Heijden P.G., Petróczi, A. Functionality of the Crosswise Model for assessing sensitive or transgressive behavior: a systematic review and meta-analysis. *Front Psychol* **2021**, *12*, 655592. <https://doi.org/10.3389/fpsyg.2021.655592>
97. Ferrin, D. L.; Gillespie, N. Trust differences across national-societal cultures: Much to do, or much ado about nothing. In Saunders M.N.K, Skinner, D., Diets, G., Gillespie, N., Lewicki, R.J. (Eds) *Organizational trust: A cultural perspective*, pp 42-86. Cambridge University Press, **2010**.
98. Nauright, J.; Ratcliff, L.; Zipp, S. Beyond scapegoats: Doping and the myth of the level playing field. *Perform. Enhanc. Health* **2025**, *13*(4), 100376. <https://doi.org/10.1016/j.peh.2025.100376>

99. Nauright, J.; Ratcliff, L.; Zipp, S. Corrigendum to "Beyond scapegoats: Doping and the myth of the level playing field". *Perform. Enhanc. Health* **2026**, *13*, 100376. <https://doi.org/10.1016/j.peh.2026.100414>
100. Ulrich, R.; Cléret, L.; Kanayama, G.; Simon, P.; Pope Jr, H. G. Clarification regarding doping rates in the article by Nauright, Ratcliff & Zipp (2025). *Perform. Enhanc. Health* **2026**, *14*(2), 100415. <https://doi.org/10.1016/j.peh.2026.100415>
101. Backhouse, S. H.; Patterson, L. B. Bridging research and practice in the psychology of doping in sport: Reflections and future directions. *Psychol. Sport Exerc.* **2026**, 103033. <https://doi.org/10.1016/j.psychsport.2025.103033>
102. Petróczi, A.; Blank, C. Progress or performance in anti-doping (social science) research? A critical reflection on achievements and future directions. *Psychol. Sport Exerc.* **2026**, 103052. <https://doi.org/10.1016/j.psychsport.2025.103052>
103. Bartels, J.M.; Schoenrade, P. The implicit association test in introductory psychology textbooks: blind spot for controversy. *Psychol Learn Teach* **2022**, *21*, 113–125.
104. Ratliff, K.A.; Smith, C.T. The implicit association test. *Daedalus* **2024**, *153*, 51–64. https://doi.org/10.1162/daed_a_02048
105. Schimmack U. The Implicit Association Test: a method in search of a construct. *Perspect Psychol Sci* **2021**, *16*, 396–414. <https://doi.org/10.1177/1745691619863798>
106. Tahamata, V.M.; Tseng, P. What does the implicit association test really measure? Insights from the theoretical debate. *Psychologia* **2024**, *66*, 137–148. <https://doi.org/10.2117/psysoc.2023-A247>
107. Barrada, J.R.; Meule, A. Orthorexia nervosa: research based on invalid measures is invalid. *J Global Health* **2024**, *14*, 03007. <https://doi.org/10.7189/jogh.14.03007>
108. Ng, Q.X.; Lee, D.Y.; Yau, C.E.; Han, M.X.; Liew, J.J.; Teoh, S.E.; Ong, C.; Yaow, C.Y.; Chee, K.T. On orthorexia nervosa: a systematic review of reviews. *Psychopathology* **2024**, *57*, 345–358. <https://doi.org/10.1159/000536379>
109. Landsheer, J.A.; Van Der Heijden, P.; Van Gils, G. Trust and understanding, two psychological aspects of randomized response. *Qual Quant* **1999**, *33*, 1–12. <https://doi.org/10.1023/A:1004361819974>
110. Lensvelt-Mulders, G.J.; Boeije, H.R. Evaluating compliance with a computer assisted randomized response technique: a qualitative study into the origins of lying and cheating. *Comput Hum Behav* **2007**, *23*, 591–608. <https://doi.org/10.1016/j.chb.2004.11.001>
111. Zuccala, A. Modeling the invisible college. *J Am Soc Inf Sci Technol* **2006**, *5*, 152–68. <https://doi.org/10.1002/asi.20256>
112. Kiss, A.; Lakner, Z.; Soós, S.; Petróczi, A. Women's footprint in anti-doping sciences: A bibliometric approach to research impact. *Front. Sports Act. Living* **2022**, *4*, 866648. <https://doi.org/10.3389/fspor.2022.866648>
113. Petróczi A, Nolte K, Schneider AJA. Women in anti-doping sciences & integrity in sport: 2021/22. *Front. Sports Act. Living* **2023**, *5*, 1248720. <https://doi.org/10.3389/fspor.2023.1248720>
114. Johnston, M. P. Secondary data analysis: A method of which the time has come. *Qual. Quant. Methods Libr.* **2014**, *3*(3), 619-626.
115. Lishner, D. A. HARKing: Conceptualizations, harms, and two fundamental remedies. *J. Theor. Philos. Psychol.* **2021**, *41*(4), 248. <https://doi.org/10.1037/teo0000182>
116. Rubin, M. When does HARKing hurt? Identifying when different types of undisclosed post hoc hypothesizing harm scientific progress. *Rev. Gen. Psychol.* **2017**, *21*(4), 308-320. <https://doi.org/10.1037/gpr0000128>
117. Baldwin, J.R.; Pingault, J.B.; Schoeler, T.; Sallis, H. M.; Munafò, M. R. Protecting against researcher bias in secondary data analysis: challenges and potential solutions. *Eur. J Epidemiol*, **2022**, *37*, 1–10. <https://doi.org/10.1007/s10654-021-00839-0>
118. Weston, S.J.; Ritchie, S.J.; Rohrer, J.M.; Przybylski, A.K. Recommendations for increasing the transparency of analysis of preexisting data sets. *Adv. Methods Pract. Psychol. Sci.* **2019**, *2*(3), 214-227. <https://doi.org/10.1177/2515245919848684>

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.