Journal of Statistics & Management Systems
ISSN 0972-0510 (Print), ISSN 2169-0014 (Online)
Vol. 29 (2026), No. 4, pp. 431–442
DOI : 10.47974/JSMS-1590

Under reporting of age by collegiate student athletes in India : A prevalence study using a mixture binary RRT model

Geeta Kalucha *
Department of Mathematics
P. G. D. A. V. College
University of Delhi
New Delhi 110065
Delhi
India

Vaani Kalucha †
Step By Step School
Noida 201304
Uttar Pradesh
India

Abstract

Age misrepresentation fundamentally undermines fairness and resource allocation within Indian sports. This study applies a Mixture Binary Randomized Response Technique (RRT) among university athletes at the campus of a Delhi University college to estimate the true prevalence of age fraud while assessing trust in the method's privacy guarantee. Using the direct-question approach anonymously, prevalence was found to be 0.300 (variance 0.000105), trust in RRT's privacy protection was 0.750 (variance 0.002526), and the RRT- based estimated prevalence was 0.260606 (variance 0.005615). These findings reveal substantial sensitivity surrounding age fraud and strong acceptability of RRT's privacy guarantees. The methodological advantages, practical implementation guidance, and policy relevance for integrating RRT into national anti-age fraud frameworks are discussed in depth.

Subject Classification: *62D05.*

Keywords: *Age fraud in sports, Randomized response technique (RRT), Respondent privacy, Social desirability bias (SDB), Trust in RRT, Variance.*

* *E-mail:* **geeta.kalucha@pgdav.du.ac.in (Corresponding Author)**
† *E-mail:* **vaanikalucha19@gmail.com**

1. Introduction and Context

Age misrepresentation involves deliberate falsification of chronological age to secure competitive advantages and has persisted in Indian sports for decades. This issue has affected multiple disciplines, including cricket, badminton, athletics, football, and wrestling. Prominent cases include cricketer Rasikh Salam's two-year ban by the Board of Control for Cricket in India (BCCI) after age-related discrepancies were discovered, legal proceedings against badminton players Lakshya Sen and Chirag Sen, and the suspension of Manjot Kalra by the Delhi Cricket Association [2]. There are other undocumented cases abound across national and regional levels. Root causes include uneven civil registration quality, powerful scholarship incentives creating motivation for fraud, and inconsistent enforcement protocols across federations [9].

Traditional detection methods like radiographic age estimation, document verification, and direct questioning face limitations such as variable accuracy, ethical concerns, and social desirability bias, impeding reliable prevalence measurement [16]. Systematic reviews have shown that the validity and reproducibility of bone/dental maturity indices for chronological age estimation using radiographic methods can vary considerably depending on population characteristics and methodology [11]. These constraints impede accurate prevalence measurements, emphasizing the need for robust, privacy-sensitive methods.

Despite the adoption of the National Code Against Age Fraud in Sports (NCAAFS) [14] and Tanner-Whitehouse 3 (TW3) radiographic verification protocols accurate quantification of age fraud prevalence remains challenging and elusive due to methodological and social barriers [2],[14]. India's civil registration system, governed by the Registration of Births and Deaths Act 1969 [17] was historically decentralized with uneven implementation and rural coverage deficits. A 2023 amendment introduced mandatory centralized, digital birth registration through the Civil Registration System portal, standardizing birth certificates as official age proof for all purposes including sports verification in an effort to reduce age fraud opportunities [12].

2. Drivers of Age Fraud and Regulatory Landscape

Age fraud in Indian sports is driven by issues in civil documentation, incentive structures emphasizing early talent categorization, and lax enforcement. Enforcement remains fragmented, with different state associations applying diverse standards and verification rigor. Additionally, cultural norms in India, where age may be perceived less rigidly than in western societies, combined with the availability of multiple document types (birth certificates, school leaving certificates, Aadhaar cards), facilitate a practice known as "age shopping," whereby athletes or their agents select the most favourable document to inflate or deflate their official age [9]. Internationally, International Federation of Association Football (FIFA) recognizes age fraud as common in youth football, a problem echoed in Indian athletics.

The regulatory framework has seen important developments. The National Code Against Age Fraud in Sports (NCAAFS), first promulgated in 2010, was comprehensively revised in 2025 to introduce ten key provisions including mandatory age verification, digital document locking to prevent tampering, and enhanced penalties for violations. These provisions apply broadly to National Sports Federations, Sports Authority of India (SAI) programs, and state associations [14]. The Board of Control for Cricket (BCCI)'s initiatives include instituting voluntary disclosure schemes [7], employing TW3 radiographic testing administered by empanelled radiologists, setting up dedicated investigation units, and integrating Aadhaar-based digital verification. The Sports Authority of India (SAI) employs a combination of TW3 and Greulich-Pyle radiographic methods, conducting verification assays on over 3,000 athletes across India's major cities, while the Khelo India scheme mandates age verification for every scholarship applicant [2],[3],[7]. Delhi University's large collegiate sports program highlights the concerns about selection efficacy and age advantages amid investments [15]. Radiographic methods (TW3, Greulich-Pyle) face ethnic bias, accuracy loss past bone fusion age (~16-17 years), and verification gaps. Document verification's multiplicity and fraud risk, plus social desirability bias impacting survey responses [10], limit direct measurement reliability.

3. Research Innovation: Applying the Randomized Response Technique (RRT)

Randomized Response Technology (RRT) allows individuals to answer sensitive questions indirectly, such as whether they have ever committed a crime or used illegal substances, by using random devices like dice or coins. Prevalence measurement of sensitive behaviours requires privacy protection to reduce under-reporting. Randomized Response Technique (RRT) provides that protection by design, enabling more candid disclosure while preserving plausible deniability for respondents. This method helps protect their privacy while providing researchers with accurate estimates of the prevalence of such behaviours in the population. Operationalizing a modern RRT variant in collegiate sport fills a key evidence gap for policy and practice.

RRT's ethical, cost-effective, and scalable properties suit the Indian sports context and sensitive nature of age fraud reporting. This study introduces the Binary Mixture Randomized Response Technique (RRT) to address the above challenges by generating more reliable estimates of age misrepresentation prevalence. The RRT methodology provides respondents with privacy protection and plausible deniability to reduce evasive answers and response bias [5],[19]. This approach is ethically acceptable, avoiding invasive medical or legal demands, and cost-effective relative to radiographic testing. It also scales efficiently for large, diverse sporting populations, making it highly suitable for national-level use.

Randomized Response Technique was originally proposed by Warner in 1965 to mitigate evasive answering bias when surveying sensitive behaviours

[19]. Greenberg et al.'s [5] extension incorporated unrelated questions to further strengthen the privacy protection. Gupta, Gupta, and Singh [6] proposed an alternative RRT model known as Optional RRT model that allowed reporting scrambled responses by respondents only if the question was found sensitive by them. It was shown that the optional models are more efficient than the non-optional models both theoretically and empirically. Aloraini et al. [1] estimated the population variance of a sensitive study variable using the optional model. Motivated by a recent study by Kalucha et al. [8] on depression among college students using a complex RRT model, we use less complex user-friendly model Mixture Binary RRT model by Lovig et al. [10], which combines direct sensitive, indirect sensitive, and unrelated control questions with an explicit truthfulness parameter, balancing privacy concerns with statistical efficiency. Trust in privacy mechanisms is a practical determinant of data quality; measuring trust alongside prevalence supports realistic deployment planning. This study pioneers its use in Indian collegiate sports age fraud detection. As this study draws heavily from the Lovig et al. model [10], the model is briefly outlined in section 3.1.

3.1 *Lovig et al. [10] – Accounting for Untruthful Responding*

The Lovig et al. [10] mixture model combines routes that include a direct sensitive question, an indirect randomized route, and an unrelated innocuous question, chosen with known probabilities. The model combined the elements of the Greenberg Unrelated Question model and the Warner Indirect Question model. It investigated the impact of untruthfulness on binary RRT models. It was seen not accounting for untruthfulness can lead to inaccurate estimates Only a fraction A of respondents facing the sensitive question will answer truthfully. The observed "yes" rate is thus a mixture of contributions from each route and the unrelated-question probability.

The following notations are used throughout the model:

$\pi_x =$ Proportion of the sensitive characteristic within the population being studied.

$\pi_y =$ Proportion of the unrelated characteristic within the population being studied.

$P_y =$ Probability of a "yes" response.

$p =$ Probability of a respondent being assigned to the direct question group.

$q =$ probability of a respondent being assigned to the indirect question group.

$1 - p - q =$ Probability of a respondent being assigned to the unrelated question group.

$A =$ Proportion of people truthfully answering to a sensitive question.

$1 - A =$ Proportion of people not responding truthfully to a sensitive question

$n =$ random sample size, with replacement.

According to this model, the probability of a "yes"response P_y is given by under this model and the corresponding estimator of π_x are given by

$$P_y = \pi_x A(p - q) + q + \pi_y(1 - p - q)\pi_y \quad \text{(i)}$$

The estimator of π_x is given by

$$\widehat{\pi_x} = \frac{\widehat{P_y} - q - (1-p-q)\pi_y}{\hat{A}(p-q)}; p \neq q \quad \text{(ii)}$$

The above estimator of π_x requires estimation of the trust level A. The estimator of the trust probability A is obtained by using the Greenberg model. The estimator and the corresponding estimated variance are given by

$$\hat{A} = \frac{\widehat{P_{y0}} - (1-p_0)\pi_{y0}}{p_0} \quad \text{(iii)}$$

The corresponding estimated variance of the above estimator is given by

$$\widehat{\text{Var}}\left[\hat{A}\right] = \frac{\widehat{P_{y0}}\left(1-\widehat{P_{y0}}\right)}{np_0^2} \quad \text{(iv)}$$

$\widehat{P_{y0}}$ = proportion of "yes"responses within the total sample of Greenberg responses;

p_0 = direct question proportion used for estimating trust/truthfulness in the Greenberg model;

π_{yg} = proportion of the unrelated characteristic in the Greenberg model.

Further,

$$\widehat{\text{Var}}[\widehat{\pi_x}] = \left[\frac{\widehat{P_y} - q - (1-p-q)\pi_y}{\hat{A}^2(p-q)}\right]^2 \widehat{\text{Var}}\left[\hat{A}\right] + \left[\frac{1}{A\widehat{(p-q)}}\right]^2 \widehat{\text{Var}}\left[\hat{P}_y\right];\ p + q \neq 1 \quad \text{(v)}$$

where the sum of the probabilities p and q is not equal to 1. The parameters in the above equation are given by

$\hat{P}_y =$ "yes" response rate of the rigged question in Equation (i);

$\widehat{\text{Var}}\left[\hat{A}\right]$ is as given by Equation (iv) , and

$$\widehat{\text{Var}}\left[\hat{P}_y\right] = \frac{\widehat{P_y}\left(1-\widehat{P_y}\right)}{n}.$$

This framework explicitly incorporates truthfulness A, enabling practitioners to diagnose and mitigate noncompliance risk while retaining estimator efficiency.

3.2 *Data Collection Protocol*

A cross-sectional field validation study was conducted at a college of University of Delhi. A sample of 200 student-athletes participated voluntarily after they were briefed about RRT sample collection survey. Participation was voluntary with informed consent; no personal identifiers were collected. The protocol minimized risk by avoiding medical tests and punitive document audits. Sample reflected the diversity of Indian collegiate athletics.

Design parameters used in the study are as recommended in Lovig et al. [10] namely,

$$p = 0.70, q = 0.15 \text{ and } \pi_y = 1/12.$$

Participants advanced sequentially through three independently operated data collection stations, with complete anonymity maintained throughout:

Data Collection Method 1: Direct anonymous questioning (Measuring the non-RRT prevalence of age misrepresentation in Sports using the Direct Drop in the Box method)

The participating students take out a slip from Box 1, check-off "yes" or "no" privately, and drop the slip in a second box. The slips had the following text:

"*Have you ever misrepresented your age for being eligible to a certain age category for a Sports event*?"

Answer: ("yes"/"no"): ------------

Data Collection Method 2: RRT Trust assessment (estimate A) using the Greenberg et al. [5] model.

Each participant drew a slip from a box consisting of 200 slips in which 140 slips (70%) presented the direct question *"Do you trust the RRT Methodology to preserve your privacy?"* and 60 slips (30%) asking the unrelated question *"Were you born in the month of April?"*. The participant drew a slip, viewed it privately, and verbally stated "yes" or "no" to the researcher, who recorded the response. The slip was then returned to the box, which was thoroughly shaken to maintain simple random selection in subsequent draws.

Data Collection Method 3: Binary Mixture RRT (Lovig et al. [10]) model implementation with direct, indirect, and unrelated control questions, measuring the prevalence of age misrepresentation (π_x).

Each participant drew a slip from a box that had three types of slips. 140 slips (70%) had the direct question "*Have you ever misrepresented your age for being eligible to a certain age category for a Sports event*?" 30 slips (15%) had the indirect question "*Have you never misrepresented your age to qualify for a specific age category for a Sports event"* and 30 slips (15%) had the unrelated question *"Were you born in the month of April?"* The participant drew a slip, privately reviewed the question, and gave a "yes" or "no" response, which the researcher recorded. The slip was placed back in the box and the box was shaken thoroughly before the next participant drew their slip to maintain simple random selection and prevent selection bias.

Regarding data quality, there were no instances of missing or invalid responses. All three clearly labelled response boxes were transported to the researcher office. The tallied responses in each box showed that a complete response rate of 100% was achieved, and it was confirmed that each container held exactly 200 responses.

Method-Specific Response Distributions:

Procedure	Method	Total no. of Yes Responses
Direct Non-RRT Prevalence Estimation	Anonymous	60
Trust Estimation	Greenberg	110
Prevalence Estimation with RRT	Lovig Mixture	49

Method 1: Non-RRT estimate of prevalence of misrepresentation of age for being eligible to a certain age category for a Sports event using direct anonymous drop- in- the- box method:

$$\widehat{\pi_x} = \frac{y}{n} = 0.3$$

$$\widehat{\text{Var}}[\widehat{\pi_x}] = \frac{\frac{y}{n}\left(1 - \frac{y}{n}\right)}{n} = 0.000105$$

where the parameter sample size$(n) = 200$ and the response rate$(y/n) = 60/200 = 0.3$.

The total count of "yes" responses (y) at the Station 1 was found to be 60.

The standard error of the estimator is

$$\text{SE}(\hat{\pi}_x) \approx 0.010247;$$

and the 95% CI are given by

$$95\%\text{ CI} \approx [0.3 \pm 1.96 \cdot 0.010247] \approx [0.2799, 0.3201]$$

Method 2: Measuring the Trust level (A):

$$\hat{A} = \frac{\widehat{P_{yg}} - (1-p_g)\pi_{yg}}{p_g} = 0.75$$

$$\widehat{\text{Var}}\left[\hat{A}\right] = \frac{\widehat{P_{yg}}\left(1-\widehat{P_{yg}}\right)}{np_g^2} \approx 0.002526$$

where:

$$\widehat{P_{yg}} = \frac{110}{200};\ \pi_{yg} = \frac{1}{12};\ p_g = 0.7$$

The standard error of the estimator is

$$\text{SE}\left(\hat{A}\right) \approx 0.050255,$$

and the 95% CI are given by

$$95\%\text{ CI} \approx [0.75 \pm 1.96 \cdot 0.050255] \approx [0.6515, 0.8485]$$

Method 3: Measuring the prevalence misrepresentation of age for being eligible to a certain age category for a Sports event (π_x)

$$\widehat{\pi_x} = \frac{\widehat{P_y} - q - \pi_y(1-p-q)}{\hat{A}(p-q)} \approx 0.260606$$

$$\widehat{\text{Var}}[\widehat{\pi_x}] \approx \left(\frac{\widehat{P_y}-q-(1-p-q)\pi_y}{\hat{A}^2(p-q)}\right)^2 \widehat{\text{Var}}[\hat{A}] + \left[\frac{1}{\hat{A}(p-q)}\right]^2 \widehat{\text{Var}}[\hat{P}_y]$$
$$\approx 0.005615$$

where:

$$\widehat{P_y} = \frac{49}{200};\ \pi_y = \frac{1}{12}$$

Consequently, the proportion 0.245 (i.e. 49/200) of "*yes*" responses resulted into a sports age misrepresentation estimate of .260606. The estimated variance turned out to be .005615.

The standard error of the estimator is

$$\text{SE}(\hat{\pi}_x) \approx 0.074934,$$

and the 95% CI are given by

$$95\%\ \text{CI} \approx [0.260606 \pm 1.96 \cdot 0.74934] \approx [0.1137, 0.4075]$$

A comprehensive comparison of the prevalence estimates, calculated variances, and 95% confidence intervals for direct anonymous questioning, trust assessment, and the Mixture Binary RRT model is presented in Table I. The table summarizes the statistical outcomes for the three primary metrics analyzed in the research:

Summary Table:

Table 1

Comparative Estimates of Age Misrepresentation Prevalence and Trust Levels across Methodological Frameworks

	METHOD 1 Non-RRT Prevalence π_x	**METHOD 2 Trust level, A**	**METHOD 3 RRT prevalence π_x**
Estimate	0.3	0.75	0.260606
Estimated Variance	0.000105	0.002526	0.005615
95% CI	[0.2799,0.3201]	[0.6515,0.8485]	[0.1137,0.4075]

4. Results and Trust in Privacy Mechanisms

We verified that if the untruthfulness is not accounted for, then the estimator for the sensitive trait is obtained by taking $A = 1$, in Equation 2 and is given by

$$\widehat{\pi_x} = \frac{\widehat{P_y}-q-\pi_y(1-p-q)}{(p-q)}$$

resulting into

$$\widehat{\pi_x} = 0.15$$

It is evident that the estimate for $\pi_x = 0.15$ accounting for trust loss to be 0.26. Notably, if untruthfulness is not accounted for, the prevalence estimate drops sharply to 15%. This underlines the critical role of modelling untruthful responses, common in sensitive contexts like age fraud reporting. The trust level of 75% indicates a substantial proportion of respondents believe in the privacy safeguards, facilitating more truthful disclosures. Statistically, the RRT's explicit truthfulness parameter differentiates between truthful denials and evasive falsehoods, allowing adjustment of prevalence estimates to reflect the actual magnitude of age misrepresentation [10]. The sizable trust parameter suggests strong acceptability of the RRT procedure, a key prerequisite for honest responding.

In sum, ignoring untruthfulness leads to an underestimated prevalence. The observed trust level corroborates that the privacy protections designed in the Mixture Binary RRT effectively encourage honest reporting, making the prevalence of 30% a realistic portrayal of this sensitive behaviour among Indian university athletes. The sizable trust parameter suggests strong acceptability of the RRT procedure, a key prerequisite for honest responding. Also, high trust levels indicate cultural acceptance of RRT methodology in Indian contexts. RRT offers benefits compared to TW3/Greulich-Pyle methods: no ethnic bias, no radiation exposure, cost-effective scalability, ethical acceptability and immediate results.

This study can motivate policy implications. The RRT prevalence estimates can guide targeted enforcement efforts toward high-risk populations, sports, and geographic regions. Annual RRT surveys could provide trend data for evaluating The National Code Against Age Fraud in Sports (NCAAFS) effectiveness. RRT could replace expensive radiographic testing for preliminary screening. The mixture design's flexibility and explicit truthfulness parameter allow tailoring to setting-specific trust and literacy constraints. Despite higher variance than direct questioning, the reduction in response bias for sensitive behaviours is the critical trade-off. Measuring A alongside prevalence makes the method self-diagnostic; if A is low, future iterations can adjust probabilities or respondent training. Prevalence evidence supports targeted compliance education and verification resource allocation where risk is higher. As with any self-report method, residual noncompliance and misunderstanding can persist; the A parameter helps monitor, but not fully eliminate, this risk. RRT can be embedded into federation audits, scholarship screening, and age-group event verification as an ethical, scalable monitoring layer.

5. Comparative International Studies and Discussion

A recent study conducted in Rivers State, Nigeria, which faces socio-economic and regulatory challenges comparable to India, reported that 74.5% of participants had encountered age fraud in sports with football (68.3%) and athletics (54.2%) being the most affected disciplines. The study identified the primary drivers of age falsification as competitive advantage (30.1%), extending playing careers (24.5%), and securing scholarships (15.7%). Furthermore, 92.5% of participants attributed poor birth registration and documentation as a major cause of age fraud, mirroring the systemic weaknesses observed in Indian sports [13]. However, this Nigerian study employed direct survey methods, and the high prevalence observed may reflect a combination of actual rates and potential over-reporting due to lack of privacy protection mechanisms. In contrast, our RRT-based prevalence estimate of 26% is lower than the drop in the box method finding of 30% but will be higher than official audit and disciplinary statistics in Indian sports, which frequently underreport the true extent of age fraud due to enforcement gaps and social desirability bias.

Globally, age fraud is widely recognized as a significant problem in youth sports, particularly in football and athletics, where verifying age through documentation remains problematic. Organizations such as International Federation of Association Football (FIFA) report frequent instances of overage players in youth competitions, with prevalence estimates ranging between 10% and 30%, varying by region and detection methodology [4]. Indirect survey techniques like RRT and biometric verification consistently yield lower but more credible prevalence estimates compared to direct questioning [16]. Recent literature also emphasizes that direct surveys, especially those conducted online or with incentives, are vulnerable to fraudulent responses and systematic under-reporting [18]. In this regard, privacy-protecting indirect methods such as RRT, particularly when coupled with high trust levels in data confidentiality, offer more precise and ethically robust prevalence measurement. Our study's trust parameter ($A = 0.75$) substantiates the reliability of RRT in eliciting truthful reporting within the Indian collegiate sports context [10].

6. Conclusion

RRT offers a robust alternative to overcome ethical and methodological limitations of existing age fraud detection methods, balancing privacy, scalability, and data quality. Integrating RRT into national frameworks like National Code Against Age Fraud in Sports (NCAAFS) enhances objective prevalence measurement vital for fair sports governance. Effective policy and education strategies depend on reliable data: a goal advanced by this privacy-protected methodology for Indian sports contexts.

7. Conflict of interest

The authors declare no conflicts of interest.

8. Acknowledgements

The authors gratefully acknowledge the invaluable mentorship and guidance of **Prof. Sat Gupta** (**University of North Carolina Greensboro,**) throughout the conception, execution, and reporting of this research. His expert advice significantly enhanced the study's design and analytical rigor.

References

[1] B. Aloraini, S. Khalil, N. Qureshi, and S. Gupta, “Variance estimation using an optional randomized response technique,” *J. Interdisciplinary Math.*, vol. 28, no. 2, pp. 369–380 (2025), doi: 10.47974/JIM-1290.

[2] BCCI, “BCCI hands Rasikh Salam two-year ban for age fraud,” ESPN Cricinfo, Jun. 18 (2019). [Online]. Available: https://sports.ndtv.com/cricket/mumbai-indians-pacer-rasikh-salam-suspended-for-two-years-for-age-fraud-2056062

[3] BCCI, “BCCI adopts additional measures to tackle age and domicile fraud in cricket,” BCCI Official Website, Sep. 25 (2020). [Online]. Available: https://www.bcci.tv/articles/2020/news/147336/bcci-adopts-additional-measures-to-tackle-age-and-domicile-fraud-in-cricket

[4] M. Frenger, E. Emrich, and W. Pitsch, “Corruption in Olympic sports: Prevalence estimations of match-fixing among German squad athletes,” *SAGE Open*, vol. 9, no. 2 (2019), doi: 10.1177/2158244019865361.

[5] B. G. Greenberg, A. L. Abul-Ela, W. R. Simmons, and D. G. Horvitz, “The unrelated question randomized response model: Theoretical framework,” *J. Amer. Statist. Assoc.*, vol. 64, no. 326, pp. 520-539 (1969).

[6] S. Gupta, B. Gupta, and S. Singh, “Estimation of sensitivity level of personal interview survey questions,” *J. Statist. Plan. Inference*, vol. 100, no. 2, pp. 239-247 (2002).

[7] Indian Express Staff, “Voluntarily declare age-fraud or face two-year ban: BCCI,” Indian Express, Aug. 2 (2020). [Online]. Available: https://indianexpress.com/article/sports/cricket/age-fraud-in-cricket-bcci-offers-amnesty-to-players-or-face-two-year-ban-6537184/

[8] G. Kalucha, S. Khalil, M. Gupta, and S. Gupta, “Estimating the Prevalence of Depression Among College Students Field Validation of a Complex Randomized Response Technique,” *J. Statist. Theory Pract.*, vol. 19, no. 96 (2025), doi: 10.1007/s42519-025-00511-w.

[9] LawInSport, “More than just a number: How India is trying to tackle age fraud in sports,” Mar. 31 (2025). [Online]. Available: https://www.lawinsport.com/topics/item/more-than-just-a-number-how-india-is-trying-to-tackle-age-fraud-in-sports

[10] M. Lovig, S. Khalil, S. Rahman, P. Sapra, and S. Gupta, “A mixture binary RRT model with a unified measure of privacy and efficien-

cy," *Commun. Statist. – Simulat. Comput.*, vol. 52, no. 6, pp. 2727–2737 (2023).

[11] V. Marconi, M. Iommi, C. Monachesi, A. Faragalli, E. Skrami, R. Gesuita, L. Ferrante and F. Carle, "Validity of age estimation methods and reproducibility of bone/dental maturity indices for chronological age estimation: a systematic review and meta-analysis of validation studies," Sci. Rep., vol. 12, Art. no. 15607, pp. 1-14 (2022).

[12] NLC Bharat, "Digitization of Birth and Death Registration," (2023). [Online]. Available: https://nlcbharat.org/digitization-of-birth-and-death-registration/

[13] D. C. Ochor, U. Bassey, and O. T. Tammy, "Age fraud in sports: Implications for sustainable talent development and athlete well-being in Rivers State," *Afr. J. Sports Phys. Sci.*, vol. 3, no. 1, pp. 25-42 (2025), doi: 10.62154/ajsps.2025.03.010685.

[14] Press Information Bureau (PIB), "Draft National Code Addresses Age Fraud, Protects Genuine Athletes," Government of India. [Online]. Available: https://www.pib.gov.in/PressReleasePage.aspx?PRID=2111175

[15] Shiksha.com, "How to get DU admissions under sports quota?" [Online]. Available: https://www.shiksha.com/humanities-social-sciences/articles/how-to-get-du-admissions-under-sports-quota-blogId-17441

[16] R. Tourangeau and T. Yan, "Sensitive questions in surveys," *Psychol. Bull.,* vol. 133, no. 5, pp. 859-883 (2007), doi: 10.1037/0033-2909.133.5.859.

[17] United Nations Statistics Division, "The Registration of Births and Deaths Act (1969)." [Online]. Available: https://unstats.un.org/unsd/dnss/docViewer.aspx?docID=1733

[18] L. Vanwersch, A. Willem, B. Constandt, and W. Hardyns, "A scoping review of the causes and consequences of fraud in Sports," J. Sport Social Issues, vol. 46, no. 6, pp. 546-584, 2022, doi: 10.1177/01937235221119811.

[19] S. L. Warner, "Randomized response: A survey technique for eliminating evasive answer bias," *J. Amer. Statist. Assoc.*, vol. 60, no. 309, pp. 63-69 (1965).

Received October, 2025