Review Article

HIDDEN ERRORS IN IVF: FROM REPORTED INCIDENTS TO ELECTRONIC WITNESSING AND PROACTIVE RISK DETECTION

***ELENA MILNER**, PhD* [1]

[1] *Independent Clinical Embryology Consultant, Haifa, Israel*

ABSTRACT

Background: Errors involving gametes and embryos are uncommon but can have irreversible clinical, psychological, ethical, and legal consequences. Conventional incident reporting captures only events that are recognized and reported, whereas electronic witnessing systems (EWS) can also reveal mismatch conditions and near misses during routine laboratory work. To review the spectrum of reported IVF laboratory errors and non-conformances, summarize published data generated by IVF laboratories using electronic witnessing and structured risk-analysis systems, and examine how automation can shift safety practice from retrospective incident review toward prospective risk detection.

Methods: Narrative review of peer-reviewed studies, contemporary professional guidance, and regulatory data addressing IVF laboratory non-conformances, traceability, witnessing, mismatch events, and electronic witnessing. No unpublished or institution-specific data are included.

Results: In a 12-year Boston IVF series, 360 graded andrology/embryology non-conformances were identified across 36,654 cycles and at least 181,899 laboratory procedures; moderate and significant non-conformances together occurred in 0.23% of cycles and 0.045% of procedures, with no major events. A 10-year Brussels IVF EWS analysis included 109,655 cycles and 849,650 witnessing points: 2,132 mismatches (0.251% per witnessing point), including 144 critical mismatches, with a yearly mean critical mismatch rate of 0.017%. At Fertility North, 73,719 witnessing points from 11,210 cycles yielded 138 mismatches (0.19%) and 16 critical mismatches (0.02%); risk varied by time of day, procedure, and operator. FMEA studies show that structured traceability and EWS implementation can substantially reduce modeled mismatch risk. UK HFEA data for 2024/25 recorded 792 reported incidents across more than 100,000 cycles and 68 near misses, while more than 99% of cycles were completed without a reported incident.

Conclusion: Serious IVF identification errors are rare, but measurable upstream mismatch conditions persist even in mature laboratories. EWS should be viewed not merely as a substitute for manual double witnessing but as a real-time safety barrier, a source of near-miss intelligence and a component of a broader quality-management system.

CITATION: *Milner E. Hidden errors in IVF: From reported incidents to electronic witnessing and proactive risk detection. Medical Times. Georgian-German Reproduction Center; 2026;(5):93–100. https://doi.org/10.71419/mtggrc.2026.45*

Keywords: IVF; electronic witnessing; near miss; specimen identification; traceability; patient safety; laboratory error; risk management; FMEA

Introduction

Assisted reproductive technology is unusually vulnerable to identification error because the biological materials being handled – oocytes, sperm, and embryos – cannot be reliably distinguished by visual inspection. At the same time, a contemporary IVF cycle requires repeated transfers between containers, workstations, and storage devices, often under substantial time pressure and cognitive load. A single failure of identification or traceability can therefore have consequences that are disproportionate to the apparent simplicity of the initiating deviation.

The published safety literature shows an important paradox. Catastrophic events such as confirmed gamete or embryo mix-up are rare, yet minor deviations, near misses and mismatch conditions are measurable. Traditional incident-reporting systems predominantly describe events that have already been recognized. Electronic witnessing can expose an additional layer of risk by recording incompatible identifiers, workflow deviations and critical mismatch conditions at the time of handling.

This review examines IVF error through three complementary evidence streams: reported incidents and non-conformances; prospective risk analysis of vulnerable laboratory processes; and real-world data generated by electronic witnessing systems. The aim is not to estimate a universal IVF error rate; published denominators and definitions differ too substantially for that, but to identify reproducible safety patterns and the role of automation in preventing or intercepting error.

Methods

This article is a narrative review. Evidence was drawn from peer-reviewed studies of IVF laboratory non-conformances, traceability, FMEA and electronic witnessing; current professional guidance from ASRM/SRBT and ESHRE; and current regulatory incident data from the UK Human Fertilisation and Embryology Authority (HFEA). Priority was given to studies reporting explicit denominators such as cycles, procedures, witnessing points or mismatch events. Because terminology differs between publications, the original terms: non-conformance, mismatch, critical mismatch, protocol deviation and near miss, are retained where possible. No unpublished data from any institution are included in this version.

The Spectrum of Error in IVF

IVF safety events range from documentation and workflow deviations to loss or damage of gametes or embryos, cryostorage failures, incorrect genetic-test handling and, at the extreme, misidentification of reproductive material. ASRM/SRBT 2026 recommends grading protocol deviations from none/minimal through moderate, significant, and major, with major events including confirmed pregnancy or birth involving misidentified sperm, oocyte, or embryo and extreme events affecting multiple patients, such as storage-tank failure. The same guidance emphasizes that near misses are valuable safety signals rather than trivial events.

Legal data illustrate the disproportionate impact of laboratory errors. In a 10-year review covering 184,015 IVF cycles across practices insured by one malpractice carrier, 176 incidents led to 30 settled claims; 21 involved indemnity payments totaling 15.062 million US$, and embryology errors accounted for eight settled claims.[11] A later US legal literature review identified 53

IVF malpractice cases from 1993-2022; 24 (45.3%) involved embryology errors, and 11 (20.8%) involved preimplantation genetic testing errors. Awards in that review ranged from 4,171 to 14.975 million US$.[12]

What Conventional Non-conformance Reporting Reveals

The most detailed long-term single-laboratory analysis remains the Boston IVF series reported by Sakkas and colleagues. Between 2003 and 2015, 36,654 treatment cycles involved at least 181,899 laboratory procedures. After exclusion of quality-control statistical anomalies, 360 andrology/embryology non-conformance reports were graded: 277 none/minimal, 66 moderate and 17 significant; no major non-conformances were reported. Moderate and significant events combined occurred in 0.23% of cycles and 0.045% of procedures, approximately one per 441 cycles or one per 2,191 procedures.[1]

The distribution of events is also informative. In the Boston IVF dataset, 52% of reported non-conformances occurred within embryology; cryopreservation accounted for 31% of moderate/significant non-conformances and PGT for 11%.[1] These findings demonstrate why a single global error rate is inadequate: risk is concentrated in specific processes, and the denominator chosen – cycle, procedure, or individual handling step – substantially changes the apparent frequency.

Regulatory Incident Data: The Visible Layer

National regulatory data provide a broader but differently defined view. The HFEA reported 792 incidents in UK-licensed clinics in 2024/25, 36% more than the 581 reported in 2023/24, across more than 100,000 treatment, storage, and donation cycles. More than 99% of cycles were completed without a reported incident. There were no Grade A incidents, 224 Grade B incidents, 443 Grade C incidents, and 68 reported near misses; 57 events were awaiting classification at the reporting cut-off. The HFEA explicitly cautions that the absence of reported incidents does not automatically indicate greater safety, as reporting patterns vary between clinics.

This distinction is central to IVF safety research. Incident databases measure detected and reported events; they do not directly measure all latent errors or interrupted error pathways. An increase in lower-grade reporting can therefore reflect a stronger safety culture rather than deterioration in care.

Why Manual Witnessing Is Not Enough

Manual double witnessing has long been used to protect chain of custody, but it is itself a human process. Repetitive checking can become automatic, require interrupting a second operator, and be vulnerable to distraction, confirmation bias, and ambiguous accountability. Electronic witnessing systems were developed to add an independent digital identification layer using barcode and/or radiofrequency identification (RFID), linking patient identity to laboratory consumables and requiring verification at predefined critical steps.

Current ASRM/SRBT guidance states that critical IVF procedures must be witnessed by a second person and/or a certified EWS, including oocyte and sperm preparation, insemination/ICSI, embryo movement, biopsy, cryopreservation, warming, transfer and disposal. The guidance also requires appropriate validation, training, documentation and investigation of electronic no-matches as near misses.[6]

Evidence From Electronic Witnessing Systems

Early electronic-witnessing observations

Early studies demonstrated that electronic witnessing could quantify mismatch conditions that conventional reporting rarely captured. A 2011 RI Witness conference abstract described 3,694 witnessing steps in 420 ART cycles. Sixteen mismatches were detected (0.43%); after excluding flow-chart/configuration and proximity-related events, four were attributed to human error, corresponding to 0.11% of witnessing steps.[9] Because this evidence was published as a conference abstract rather than a full-length peer-reviewed article, it should be interpreted cautiously.

Ten years of electronic witnessing at Brussels IVF

The largest long-term published EWS dataset was reported by Sterckx and colleagues. Across 109,655 IVF/ICSI, frozen embryo transfer and IUI cycles, 724,096 RFID tags were used, and 849,650 witnessing points were registered. The system recorded 2,132 mismatch events, an overall rate of 0.251% per witnessing point and 1.944% per cycle. Of these, 144 were classified as critical mismatches. The yearly mean critical mismatch rate was 0.017 ± 0.007% per witnessing point and 0.129 ± 0.052% per cycle. In addition, 940 administrator-assigned events were recorded, including 320 critical administrator assignments.[2]

The study is particularly valuable because it separates total alerts from events judged critical, such as mislabelling or non-matching samples within one work area. Critical mismatches represented 144/2,132 (6.75%) of all mismatch events. Sperm preparation and IVF/ICSI were the procedures most prone to critical mismatch and administrator-assign events. The authors also cautioned that EWS can introduce new risks, including blind witnessing, if used incorrectly.[2]

Risk factors identified from 73,719 witnessing points

Forbrig and colleagues provided a complementary risk-factor analysis from Fertility North in Australia. Their retrospective dataset comprised 73,719 witnessing points from 11,210 consecutive treatment cycles between January 2016 and July 2023. They identified 138 mismatches (0.19%) and 16 critical mismatches involving biological material (0.02%); no actual mix-up occurred.[3]

Risk was not uniformly distributed. The overall mismatch rate was highest between 12:00 and 14:00 (0.28%), whereas the critical mismatch rate was highest after 16:00 (0.13%). Mismatch rates differed significantly across laboratory processes and operators. Embryo transfer had the highest overall mismatch rate (0.36%), while oocyte denudation had the highest critical mismatch rate (0.13%). Multivariable analysis associated oocyte denudation, embryo transfer, cryopreservation, sperm preparation, oocyte collection, and insemination with mismatch risk; critical mismatch was, on univariate analysis, associated with work after 16:00 and oocyte denudation.[3]

Prospective Risk Analysis: FMEA and Traceability

Electronic witnessing is most effective when embedded within a broader process analysis. Rienzi and colleagues used failure mode and effects analysis (FMEA) to map seven IVF process phases, 19 process steps, and 32 possible failure modes. Before EWS implementation, the highest risk-priority number (RPN) was 30; after implementation, the highest RPN was 10.4. This does not demonstrate a two-thirds reduction in observed clinical errors; it demonstrates a reduction in the modeled risk score assigned to the highest-risk failure mode.

A subsequent multicentre Italian FMEA project developed a comprehensive IVF traceability protocol, reinforcing that identification safety requires a chain of controls rather than a sin-

gle witnessing step.[5] Current ESHRE recommendations similarly place risk assessment, patient/sample identification and traceability, witnessing, validation, audits and management of non-conformities within the IVF laboratory quality-management framework.[7]

Comparing Published Electronic-Witnessing Data

Study	Setting / period	Exposure	All mismatches	Critical / true mismatches	Key observation
Thornhill et al. 2011 (conference abstract)	ART laboratory; short-term conference report	420 cycles; 3,694 witness steps	16 (0.43%)	4 human-error true mismatches (0.11%)	Early demonstration that EWS detects and classifies mismatch conditions
Sterckx et al. 2023	Brussels IVF; 2011–2021	109,655 cycles; 849,650 witnessing points	2,132 (0.251%/WP)	144; yearly mean 0.017%/WP	Large long-term dataset; only a minority of alerts were critical
Forbrig et al. 2025	Fertility North; 2016–2023	11,210 cycles; 73,719 witnessing points	138 (0.19%)	16 (0.02%)	Risk varied by time, process and operator; no actual mix-up

Direct numerical comparison between studies should be approached with caution. Electronic platforms differ in workflow architecture and alert logic, while authors use different definitions of mismatch, true mismatch, critical mismatch and administrator intervention. Nevertheless, the studies consistently show that measurable mismatch conditions persist during routine ART work and that only a subset of total alerts represents a biologically critical event.

From Error Prevention to Risk Intelligence

The most important contribution of EWS may be conceptual. A conventional witness answers a binary question at a critical moment: do these identifiers match? A mature electronic system can also create a longitudinal dataset describing where and when alerts occur, which workflow steps generate technical noise, and whether corrective actions change recurrence. This turns witnessing from a compliance task into a potential source of safety intelligence.

Such data should not be used primarily to rank or blame individual embryologists. Operator-level variation may identify training or workflow needs, but errors in complex systems usually arise from interactions among workload, process design, interruptions, staffing, equipment, software, and human behavior. ASRM/SRBT therefore recommends nonpunitive reporting, continuous surveillance, root cause analysis, and corrective action. A just culture is especially important because under-reporting destroys the very data needed for prevention.

Cryostorage: Extending the Chain of Identity

The next logical boundary for electronic witnessing is long-term cryostorage. As cryobanks expand, safety depends not only on confirming specimen identity at freezing and warming but also on maintaining reliable linkage among patient identity, the individual cryogenic container, and its physical storage location throughout the storage lifecycle.

An integrated cryostorage-management model could make each movement into, within, or out of storage a controlled, traceable event, reducing dependence on manually maintained location records and strengthening protection against misplacement or retrieval of an incorrect specimen. Commercial approaches such as CryoControl and location-identification tools such as CryoSpacer illustrate this direction, but their clinical impact should be evaluated independently rather than inferred from product functionality alone. The broader principle is a continuous chain of custody: identity, location, and action should remain traceable from specimen creation to final disposition.

Patient acceptability is also relevant. In a prospective single-centre cohort of 408 patients, 90.4% expressed significant concern about sample mix-up; use of an EWS reduced those concerns in 92.1%, and 97.1% reported high satisfaction with electronic traceability.[10] These findings address patient perception rather than laboratory error prevention, but they illustrate the importance patients place on visible traceability safeguards.

Beyond mismatch prevention, EWS can reduce the operational burden of manual double witnessing. A prospective timing study found substantial reductions in witnessing time across ICSI, Day-3 embryo assessment, fresh embryo transfer and frozen embryo transfer.[13]

Limitations of the Evidence

The literature has important limitations. There is no universal taxonomy of IVF laboratory errors, and denominators vary across cycles, procedures, tags, and witnessing points. Reporting systems are influenced by local safety culture, regulatory requirements, and willingness to report. EWS platforms differ in hardware, software logic, and workflow configuration, and a mismatch alert does not prove that an adverse event would have occurred without intervention. Much of the evidence is retrospective and single-center. Finally, some early EWS data were published only as conference abstracts. These limitations preclude a single pooled ‘IVF error rate’ and argue for standardized definitions and multicentre reporting.

Practical Implications for IVF Laboratories

Available evidence supports a layered safety model. Laboratories should maintain unambiguous labeling and chain-of-custody procedures; witness critical specimen movements; validate and audit EWS where used; classify and investigate no-matches and protocol deviations; monitor mismatch and technical-alert trends; and use root-cause analysis and prospective risk tools such as FMEA. 4-7 Electronic witnessing may also improve workflow efficiency: in a prospective comparison of 342 witnessing observations, EWS reduced total witnessing time by approximately 3.1 – to 5.2-fold across four common IVF procedures.[13] The objective is not zero reporting. A laboratory that detects, reports, and learns from near misses may be safer than one that reports none.

Conclusion

IVF laboratory errors are rare at the level of catastrophic outcome but are not absent at the level of process deviation and near miss. Conventional non-conformance reporting, regulatory

surveillance, and electronic witnessing describe different layers of the same safety problem. Published EWS datasets involving hundreds of thousands of witnessing points show that mismatch conditions can be detected during routine work, that biologically critical mismatches form only a minority of total alerts, and that risk clusters around particular processes and working conditions. Electronic witnessing should therefore be understood as more than an automated second witness. When implemented correctly within a quality-management system, it can provide real-time interception, traceability, and a measurable dataset for proactive risk reduction. The future direction is a continuous digital chain of identity and location extending from gamete collection through embryo culture, genetic testing, cryostorage, warming, transfer, and final disposition.

Declarations

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of Interest: The authors declare no competing interests.

Declarations

Ethics approval: Not applicable. This manuscript is a narrative review and contains no original patient-level or institution-specific research data.

Funding: No funding statement has been provided; this should be completed in accordance with the target journal's requirements.

Conflict of interest: To be completed by the author according to the target journal's disclosure requirements, including any relevant advisory or commercial relationships.

Data availability: Not applicable; no original dataset was generated for this review.

References:

1. Sakkas D, Barrett CB, Alper MM. Types and frequency of non-conformances in an IVF laboratory. Hum Reprod. 2018;33 (12):2196-2204. doi:10.1093/humrep/dey320.
2. Sterckx J, Wouters K, Mateizel I, et al. Electronic witnessing in the medically assisted reproduction laboratory: insights and considerations after 10 years of use. Hum Reprod. 2023;38 (8):1529-1537. doi:10.1093/humrep/dead115.
3. Forbrig M, Peirce K, Copeland C, Natalwala J, Chapple V, Liu Y. Risk factors for mismatches in the ART laboratory: an analysis of 73,719 electronic witnessing points. Reprod Biomed Online. 2025;50:104500. doi:10.1016/j.rbmo.2024.104500.
4. Rienzi L, Bariani F, Dalla Zorza M, et al. Failure mode and effects analysis of witnessing protocols for ensuring traceability during IVF. Reprod Biomed Online. 2015;31 (4):516-522. doi:10.1016/j.rbmo.2015.06.018.
5. Rienzi L, Bariani F, Dalla Zorza M, et al.; Italian Society of Embryology, Reproduction and Research (SIERR). Comprehensive protocol of traceability during IVF: the result of a multicentre failure mode and effect analysis. Hum Reprod. 2017;32 (8):1612-1620. doi:10.1093/humrep/dex144.
6. Practice Committee of the American Society for Reproductive Medicine; Practice Committee of the Society of Reproductive Biologists and Technologists. Witnessing and protocol deviations in the in vitro fertilization and andrology laboratory: a committee opinion. Fertil Steril. 2026;126 (1):38-48. doi:10.1016/j.fertnstert.2026.03.014.

7. ESHRE Good Practice in the IVF Lab Working Group. ESHRE recommendations on Good Practice in the IVF laboratory. Hum Reprod. 2026;41 (8):1245-1269. doi:10.1093/humrep/deag096.
8. Human Fertilisation and Embryology Authority. The fertility sector 2024/25. HFEA. Published November 2025; updated April 28, 2026. Accessed August 31, 2026.
9. Thornhill AR, Brunetti XO, Bird S, Bennett K, Rios LM, Taylor J. Reducing human error in IVF with electronic witnessing. Fertil Steril. 2011;96 (3 Suppl):S179. doi:10.1016/j.fertnstert.2011.07.697.
10. Forte M, Faustini F, Maggiulli R, et al. Electronic witness system in IVF–patients' perspective. J Assist Reprod Genet. 2016;33 (9):1215-1222. doi:10.1007/s10815-016-0759-4.
11. Penzias AS, Bendikson KA, Butts SF, et al. Outcomes of medical malpractice claims in assisted reproductive technology over a 10-year period from a single carrier. J Assist Reprod Genet. 2017;34 (4):459-465. doi:10.1007/s10815-017-0889-3.
12. Applebaum J, Humphries LA, Nepps ME, Berger DS, O'Neill K. Malpractice litigation surrounding in vitro fertilization in the United States: a legal literature review. Fertil Steril. 2023;119 (4):572-580. doi:10.1016/j.fertnstert.2022.12.038.
13. Holmes R, Wirka KA, Catherino AB, Hayward B, Swain JE. Comparison of electronic versus manual witnessing of procedures within the in vitro fertilization laboratory: impact on timing and efficiency. F&S Rep. 2021;2 (2):181-188. doi:10.1016/j.xfre.2021.04.006.

Table 1. Selected published evidence relevant to IVF laboratory error and electronic witnessing

Evidence source	Population / denominator	Main findings	Interpretation
Boston IVF non-conformance database	36,654 cycles; ≥181,899 procedures	277 minimal, 66 moderate, 17 significant, 0 major; moderate+significant 0.045%/procedure	Structured reporting reveals low-frequency but recurrent process failures
Brussels IVF EWS	109,655 cycles; 849,650 witnessing points	2,132 mismatches (0.251%/WP); 144 critical; mean critical rate 0.017%/WP	Long-term EWS detects a measurable hidden layer; most alerts are not critical
Fertility North EWS	11,210 cycles; 73,719 witnessing points	138 mismatches (0.19%); 16 critical (0.02%); no mix-up	Mismatch risk varies by time, process and operator
Rienzi FMEA	7 phases; 19 steps; 32 failure modes	Highest RPN decreased from 30 to 10 after EWS	EWS plus process redesign reduced modeled mismatch risk
HFEA 2024/25	>100,000 UK cycles	792 reported incidents; 68 near misses; >99% of cycles without a reported incident	Regulatory reporting captures visible events but is sensitive to reporting culture