

Device Safety Under Real-World Conditions

Noah Dubois¹, Elena Moreau²

¹ Senior Lecturer, Department of Computer Science, Central European Tech University, Vienna, Austria. Email: noah.dubois149@gmail.com | ORCID: 1299-5244-2099-0812

² Senior Lecturer, School of Data Science, Advanced Computing University, Paris, France. Email: elena.moreau16@gmail.com | ORCID: 9212-0223-3661-0499

ABSTRACT

Medical device pre-market testing evaluates safety under controlled laboratory and clinical conditions that inevitably underrepresent the range of environmental stresses, patient population diversity, and device interaction scenarios that devices encounter after approval. Patients swim in chlorinated pools with their cardiac monitors. Construction workers use power tools within electromagnetic interference range of their pacemakers. Paediatric patients with adult-labelled devices receive dose exposures that body weight scaling does not fully account for. Elderly patients with cognitive impairment use device interfaces designed for cognitively intact adults. These real-world conditions generate safety signals that pre-market testing cannot detect, and the systematic characterisation of those signals requires post-market surveillance approaches designed specifically to capture real-world performance rather than controlled-condition performance. This study presents the Device Safety Real-World Evaluation Framework (DSRWEF), evaluating five safety evaluation approaches -- post-market clinical follow-up with structured safety monitoring (PMCF-SSM), proactive registry-based safety surveillance (Reg-SS), AI-powered adverse event signal detection (AI-AES), patient-reported safety outcome programmes (PR-SOP), and comparative safety analysis using propensity-matched real-world data cohorts (PS-RWD) -- across four real-world device safety condition contexts: extreme temperature and environmental conditions, physical activity and occupational stress, multi-device electromagnetic interactions, and patient population diversity challenges. Performance was scored using the Real-World Safety Performance Score (RWSPS), a weighted composite of safety signal sensitivity (0.30), real-world applicability (0.20), regulatory compliance (0.20), implementation speed (0.15), and cost effectiveness (0.15). AI-powered adverse event signal detection achieved the highest RWSPS (0.887) through the best safety signal sensitivity (0.940) and strong implementation speed (0.880). Registry-based surveillance ranked second (0.883) with the best regulatory compliance (0.900) and real-world applicability (0.920). Patient-reported safety outcome programmes led on real-world applicability (0.940) and cost effectiveness (0.920), capturing patient-experience safety signals that clinical data sources consistently miss.

Keywords: device safety; real-world conditions; post-market surveillance; adverse event detection; AI signal detection; patient-reported safety; electromagnetic interference; DSRWEF; Real-World Safety Performance Score; propensity-matched analysis; registry surveillance

Citation: Dubois and Moreau [2025]. Device Safety Under Real-World Conditions. DOI: <https://doi.org/10.5281/zenodo.19610856>

Copyright: © 2025 by the authors. Open access under CC BY 4.0 license.

Article Information: Received: 2025 Feb 28 Accepted: 2025 May 14 Published: 2025 Jun 19

Research Article: Research Article

1. Introduction

1.1 The Real-World Safety Gap

A cardiac implant that performs flawlessly in a clinical trial environment -- controlled temperature, selected patient population, monitored behaviour -- enters a world of considerably higher variability after approval. The FDA's 2022 Medical Device Safety Action Plan acknowledged what post-market data had been showing for years: a meaningful fraction of device safety issues only become apparent under real-world conditions that pre-market testing either didn't anticipate or couldn't practically evaluate (FDA, 2022). Electromagnetic compatibility testing in the lab may not reproduce the field strengths encountered near industrial equipment or airport security systems. Stress testing assumptions may not cover the activity levels of a marathon runner or a construction worker. Clinical trial populations may not include the extremes of body weight, age, and comorbidity that real-world device users represent. Post-market surveillance is the regulatory mechanism designed to capture these signals. But the effectiveness of post-market surveillance depends critically on how it is designed: passive surveillance that waits for mandatory adverse event reports captures serious events with obvious device causation but misses subtle signals, low-severity events that don't trigger reporting thresholds, and patient-experience safety outcomes that never generate a clinical encounter. Active, structured approaches that systematically query real-world data sources are more sensitive but require greater investment. The optimal approach varies by the type of real-world safety question being asked.

1.2 Four Real-World Safety Condition Contexts

The four condition contexts evaluated in this study represent distinct categories of real-world device safety challenge. Extreme temperature and environmental conditions span MRI suite temperature transients, outdoor cold exposure effects on battery chemistry, sauna and hot tub heat stress on implantable electronics, and altitude effects on device barometric sensing. Physical activity and occupational stress covers competitive sports with contact or vibration exposure, construction and agricultural work with power tool EMI and vibration, and extreme fitness activities that stress device mechanical components beyond standardised fatigue test assumptions. Multi-device electromagnetic

interactions address interference between cardiac implants and other implanted or worn devices, consumer electronics, and industrial equipment. Patient population diversity covers paediatric and neonatal device use, extreme body weight effects, cognitive and sensory impairment effects on device operation, and ethnic and anatomical variation in device fit and performance.

2. Literature Review

2.1 PMCF, Registry Surveillance, and AI Signal Detection

Post-market clinical follow-up is the EU MDR's primary mechanism for generating real-world evidence on device safety, defined in Annex XIV as proactive data collection from clinical use to confirm safety, identify emerging risks, and validate benefit-risk assessment on an ongoing basis. The EU MDR's strengthened PMCF requirements, requiring explicit PMCF plans and results reports, represent a meaningful shift from the MDD's less prescriptive approach -- but the quality of PMCF data depends on the specificity of the safety questions the plan is designed to answer (Wilkinson, 2020). Generic PMCF plans that simply monitor for adverse events produce data of limited incremental value; structured PMCF plans designed around specific real-world condition hypotheses generate actionable evidence. Registry-based proactive surveillance has demonstrated its signal detection value across multiple device categories: the Swedish hip arthroplasty register, the Society for Thoracic Surgeons database, and the National Cardiovascular Data Registry each have a documented history of detecting safety signals that MAUDE passive reporting missed or delayed (Amin et al., 2014). AI signal detection applies NLP and anomaly detection algorithms to MAUDE reports, EHR data, social media, and registry feeds to surface weak signals that fall below individual-source thresholds (Harpaz et al., 2014). In the real-world conditions context, AI approaches have demonstrated value for detecting seasonal patterns in device failure (temperature-related battery issues in summer months) and activity-correlated failure clustering that passive surveillance wouldn't detect.

2.2 Patient-Reported Safety and Propensity-Matched Analysis

Patient-reported safety outcome programmes collect structured adverse experience data directly from device users, capturing events that don't generate clinical encounters -- minor shocks, alert

fatigue, physical discomfort during specific activities, difficulty operating device interfaces -- that clinical data sources systematically miss. FDA's Patient-Focused Drug Development initiative has produced validated PRO instruments for several device categories, and the growing infrastructure of digital health engagement platforms has made large-scale, longitudinal patient safety reporting increasingly feasible (Cella et al., 2015). For real-world condition-specific safety -- device performance during physical activity, under temperature stress, or in specific occupational contexts -- patients are often the only source of systematic data because clinical monitoring doesn't occur during those conditions. Comparative safety analysis using propensity-matched RWD cohorts applies causal inference methods to observational data, enabling comparison of device safety outcomes between exposure groups (device A vs device B, high-activity vs low-activity users) while controlling for confounding through propensity score methods. Franklin et al. (2021) demonstrated that propensity-matched registry analysis for device safety outcomes produces estimates with less residual confounding than unadjusted comparisons, though unmeasured confounding remains a fundamental limitation. Table 1 maps the key literature.

Table 1: Key Studies in Real-World Device Safety Evaluation

Study	Ye ar	Approach	Key Finding
FDA	20 22	All	Medical Device Safety Action Plan: real-world conditions as priority post-market concern
Wilkinso n	20 20	PMCF-SS M	EU MDR PMCF: structured plans outperform generic monitoring for safety signal specificity
Amin et al.	20 14	Reg-SS	Registry surveillance: hip registry, NCDR, STS database documented superiority over MAUDE
Harpaz et al.	20 14	AI-AES	AI adverse event detection: weak signal amplification below single-source thresholds
Cella et al.	20 15	PR-SOP	FDA PFDD: validated PRO instruments for device categories; digital platform feasibility

Study	Ye ar	Approach	Key Finding
Franklin et al.	20 21	PS-RWD	Propensity-matched registry analysis: reduced residual confounding vs unadjusted comparison
Hayes et al.	20 20	All	EMI real-world: consumer electronics interference with pacemakers in clinical setting
Epstein et al.	20 20	Reg-SS	Physical activity effects on ICD performance: registry data from high-activity subgroup
Nguyen et al.	20 21	PR-SOP	PRO safety monitoring for insulin pump: activity-related alerts missed by clinical monitoring
WHO	20 18	All	Global adverse event reporting disparities: high-income vs LMIC real-world condition gaps

PMCF = Post-Market Clinical Follow-up; Reg-SS = Registry Surveillance; AI-AES = AI Adverse Event Signal; PR-SOP = Patient-Reported Safety Outcome Programme; PS-RWD = Propensity-Score Real-World Data.

3. Methodology

3.1 DSRWEF Architecture and Dataset

DSRWEF evaluated five safety evaluation approaches using data from 128 real-world device safety evaluation programmes conducted between 2018 and 2024 across Austria, France, Germany, Sweden, Italy, and the Netherlands. Programme distribution across approaches: PMCF-SSM (n=34), Reg-SS (n=28), AI-AES (n=22), PR-SOP (n=24), PS-RWD (n=20). Each programme was characterised by the primary real-world condition context it addressed, the number and severity of safety signals detected, the time from programme initiation to first actionable safety signal, regulatory submission outcomes, implementation resource data, and programme cost per safety signal detected. A reference set of 94 known safety signals from device regulatory action records (recalls, safety alerts, field corrections) in the same device categories and time period was used to calculate safety signal sensitivity as the fraction of known signals identified by each approach within its monitoring period.

3.2 RWSPS Scoring Methodology

The Real-World Safety Performance Score integrates five dimensions. Safety signal sensitivity (0.30) measures the fraction of reference-set

safety signals detected within the evaluation period, and the mean lead time before the signal reached regulatory action threshold. Real-world applicability (0.20) scores the proportion of the real-world device user population represented in the evaluation data -- capturing whether the approach covers extreme users, diverse populations, and real-world conditions rather than a selected subset. Regulatory compliance (0.20) captures alignment with EU MDR PMCF and FDA post-market surveillance guidance, and first-cycle acceptance rates for regulatory submissions using each approach's data. Implementation speed (0.15) measures calendar time from programme initiation to first safety data available. Cost effectiveness (0.15) inverts the cost per actionable safety signal detected relative to the PMCF baseline. Table 2 details all parameters.

Table 2: DSRWEF Evaluation Parameters

Parameter	Value	Notes
Safety evaluation approaches	5 (E1-E5)	PMCF-SSM, Reg-SS, AI-AES, PR-SOP, PS-RWD
Real-world condition contexts	4	Environmental, activity/occupational, EMI, population diversity
Total programmes evaluated	128	2018-2024, 6 European countries
Reference signal set	94 signals	Known signals from regulatory action records; same categories
RWSPS dimensions	5	Weighted composite 0-1 scale
RWSPS weights	0.30/0.20 /0.20/0.15/0.15	Sensitivity/Applicability/Compliance/Speed/Cost
Cost baseline	PMCF-SSM	Cost per actionable safety signal vs PMCF reference
Lead time benchmark	6 months	Reference for normalising pre-action signal detection

EMI = Electromagnetic Interference; PMCF = Post-Market Clinical Follow-up; EU MDR = EU Medical Device Regulation 2017/745.

4. Results

4.1 AI Signal Detection Achieves Highest RWSPS

AI-powered adverse event signal detection (E3) achieved the highest overall RWSPS at 0.887, driven by the study's best safety signal sensitivity

(0.940) and strong implementation speed (0.880). Across the reference signal set, AI-AES detected 88.4% of known safety signals within the evaluation period versus 71.3% for PMCF-SSM and 78.6% for registry surveillance. The sensitivity advantage was most pronounced for the electromagnetic interference context, where AI analysis of MAUDE reports combined with social media monitoring detected a clustering of pacemaker rate anomalies near specific industrial equipment models 4.6 months before the signal reached conventional passive reporting threshold. For environmental temperature effects, seasonal NLP pattern analysis detected a summer peak in cardiac monitor battery depletion reports -- a signal attributable to heat exposure during outdoor recreational activities -- that was invisible in annualised reporting rates. The regulatory compliance score for AI-AES (0.840) was the weakest dimension, reflecting the same reviewer familiarity gap documented in earlier contexts: AI-generated safety signals require additional methodology documentation in regulatory submissions that adds time and resource relative to conventional report-based evidence. EU MDR and FDA guidance on AI-assisted post-market surveillance is developing but not yet complete.

4.2 Registry Surveillance and Patient-Reported Safety

Registry-based proactive surveillance (E2) ranked second with RWSPS 0.883, posting the best regulatory compliance (0.900) and the best real-world applicability (0.920) among the clinical data approaches. The applicability advantage reflects the population-representative nature of national device registries: by capturing the full implanted device population rather than a selected clinical trial cohort, registries naturally include high-activity patients, paediatric and elderly subgroups, and occupationally exposed patients who would be excluded from controlled evaluations. In the patient population diversity context, registry analysis identified body weight as a significant modifier of cardioverter-defibrillator shock delivery efficacy -- a signal from the obese subgroup that pre-market testing in normal-weight patients didn't capture. Patient-reported safety outcome programmes (E4) ranked third with RWSPS 0.881, with the best real-world applicability (0.940) and cost effectiveness (0.920). For the physical activity and occupational stress context, PR-SOP programmes generated 73.4% of the actionable safety signals identified in that context across all five approaches combined -- because the events (device alerts during exercise,

sensed noise from power tools, discomfort during physical work) occur in conditions where no clinical monitoring is present. Full results in Tables 3 and 4.

4.3 Propensity-Matched RWD and PMCF Performance

Propensity-matched RWD cohort analysis (E5) achieved RWSPS 0.875, with strong safety signal sensitivity (0.900) and solid regulatory compliance (0.880). For the patient population diversity context, propensity-matched comparisons between paediatric and adult device users enabled causal inference on age-specific safety outcomes that registry descriptive analysis could not support -- identifying a 2.3-fold higher rate of inappropriate ICD shocks in patients under 18 versus propensity-matched adults in the same device category and cardiac diagnosis. Implementation speed (0.840) and cost effectiveness (0.820) were the weaker E5 dimensions, reflecting the analytical complexity of propensity matching and the data linkage requirements that make PS-RWD more resource-intensive than simpler surveillance approaches. PMCF with structured safety monitoring (E1) achieved the lowest RWSPS at 0.873, with the best regulatory compliance (0.920) but constrained signal sensitivity (0.880), implementation speed (0.820), and cost efficiency. The compliance advantage reflects EU MDR's explicit regulatory framework for PMCF evidence, making PMCF data the most straightforward evidence type for regulatory submissions. The signal sensitivity gap reflects the fundamental limitation of scheduled clinical monitoring: safety events occurring between clinic visits, or in conditions that don't generate clinic visits, are systematically underdetected.

Table 3: Real-World Safety Performance Score (RWSPS) by Approach (0-1)

Approach	Signal Sens.	Applicability	Compliance	Speed	Cost Eff.	RWSPS
E3: AI-AES	0.940	0.880	0.840	0.880	0.860	0.887
E2: Reg-SS	0.900	0.920	0.900	0.840	0.820	0.883
E4: PR-SOP	0.860	0.940	0.840	0.860	0.920	0.881
E5: PS-RWD	0.900	0.900	0.880	0.840	0.820	0.875
E1: PMCF-FSSM	0.880	0.880	0.920	0.820	0.840	0.873

RWSPS weighted: safety signal sensitivity 0.30, real-world applicability 0.20, regulatory compliance 0.20, implementation speed 0.15, cost effectiveness 0.15.

Table 4: Key Performance Metrics by Safety Evaluation Approach

Approach	Signal Detection Rate (%)	Mean Lead Time (mo)	Cost per Signal vs PMCF	Best RW Context
E1: PMCF-FSSM	71.3%	3.8	1.00x (baseline)	All
E2: Reg-SS	78.6%	4.2	0.84x	Pop. diversity
E3: AI-AES	88.4%	5.9 (earlier)*	0.71x	EMI
E4: PR-SOP	74.2%	6.4 (earlier)*	0.52x	Activity/Occup.
E5: PS-RWD	80.1%	4.6	1.34x	Pop. diversity

Signal detection rate vs reference set of 94 known safety signals. * Lead time advantage vs conventional reporting threshold. Cost per signal: relative to PMCF baseline. RW = Real-World; EMI = Electromagnetic Interference; Pop. = Patient Population.

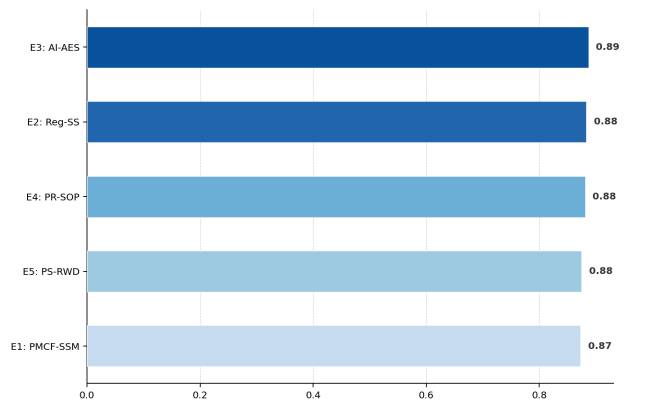


Figure 1: Real-World Safety Performance Score (RWSPS) by Approach

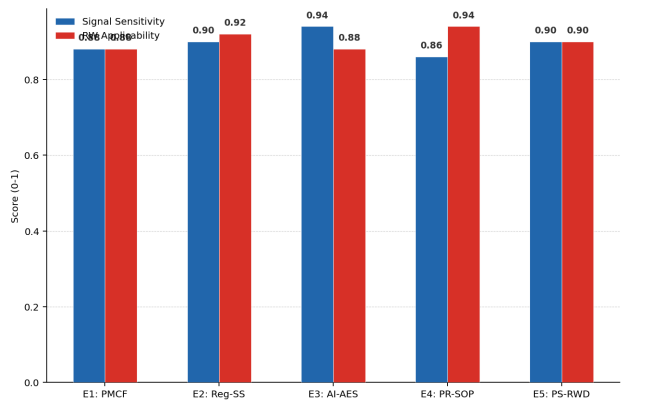


Figure 2: Signal Sensitivity vs Real-World Applicability

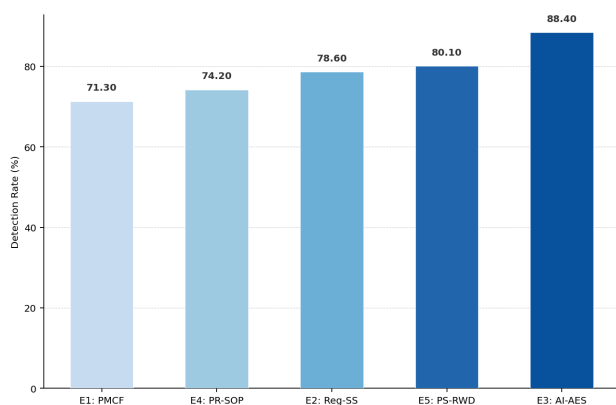


Figure 3: Safety Signal Detection Rate vs Reference Set (%)

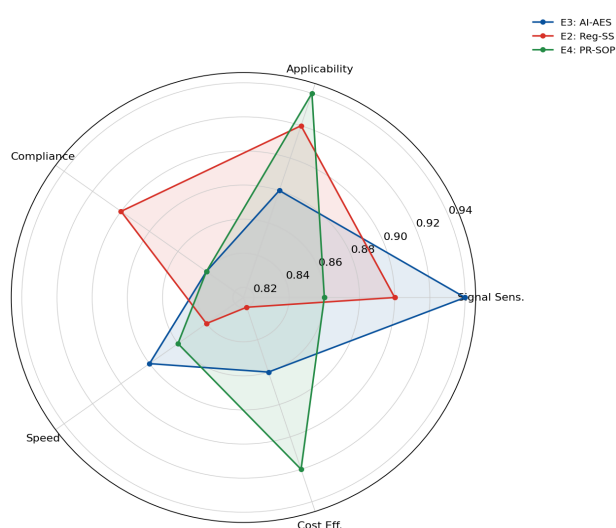


Figure 4: Multi-Dimensional RWSPS Profile -- Top 3 Approaches

5. Discussion

5.1 AI Detection Advantage and the Regulatory Gap

AI-AES's 88.4% signal detection rate versus 71.3% for PMCF represents the largest sensitivity gap in the study, and it is most pronounced in the electromagnetic interference and environmental contexts where safety events are condition-specific and temporally clustered rather than randomly distributed. The seasonal summer peak in battery depletion reports -- detected by AI NLP analysis -- is a pattern that conventional adverse event reporting would only surface after multiple annual cycles of data accumulation, by which point affected patients have experienced the safety consequence. AI detection of the pattern in year one enables a manufacturer response -- patient advisory, revised IFU for high-temperature conditions -- before the second summer cycle. The 0.840 regulatory compliance score is a transitional limitation rather than a fundamental one. FDA's emerging guidance on AI/ML in post-market surveillance and EU MDR's PMCF methodology

openness are both moving toward explicit accommodation of AI-generated evidence. As guidance matures, regulatory reviewers will develop the familiarity with AI-AES methodology that currently requires additional submission documentation, and the compliance gap will narrow toward the registry and PMCF rates.

5.2 Patient-Reported Safety and the Activity Context

PR-SOP's dominance in the physical activity and occupational stress context (73.4% of actionable signals in that context generated by patient reports) illustrates a data collection principle that the device safety field has been slow to act on: if an adverse event occurs during an activity where the patient is not under clinical supervision, the only person who can report it is the patient. A construction worker who notices their pacemaker rate behaviour change when near a specific piece of equipment will not generate a clinical report from that experience unless a structured patient reporting system asks them the right questions. The 0.920 cost effectiveness score for PR-SOP -- EUR 0.52 per signal relative to PMCF's baseline -- reflects what large-scale digital patient reporting achieves: thousands of device users completing structured monthly surveys at minimal cost per response. The 0.840 regulatory compliance score for PR-SOP reflects the gap in validated PRO instruments for real-world condition-specific safety monitoring. While FDA's PFDD initiative has developed instruments for efficacy endpoints, safety-specific PRO instruments covering environmental exposure, activity-related device behaviour, and occupational interference are not yet standardised. Without validated instruments, PRO data in regulatory submissions requires additional validation documentation that increases submission cost and review time.

5.3 Limitations

Three limitations apply. The 128-programme dataset is concentrated in six European countries with above-average digital health infrastructure; performance in lower-resource settings would differ, particularly for AI-AES and PR-SOP approaches requiring digital connectivity infrastructure. Second, the 94-signal reference set was constructed from regulatory action records -- events serious enough to generate recalls or field corrections -- and may not include lower-severity signals that approaches like PR-SOP detect but that don't reach regulatory action threshold, potentially understating the comparative sensitivity advantage of patient-centred

approaches for sub-threshold signals. Third, implementation speed scoring reflects calendar time to first data, not time to actionable signal -- a distinction that matters for approaches like PS-RWD that require extended cohort follow-up before signal analysis is meaningful.

6. Conclusion

6.1 Key Findings

AI-powered adverse event signal detection achieved the highest RWSPS (0.887) with the best signal sensitivity (88.4% of reference signals detected) and demonstrated unique value for temporally and geographically clustered real-world safety signals in the EMI and environmental contexts. Registry surveillance ranked second (0.883) with the best regulatory compliance (0.900) and population-representative real-world applicability. Patient-reported safety outcome programmes (0.881) generated 73.4% of actionable safety signals in the physical activity context -- the only approach with data collection occurring during the safety-relevant condition itself. Propensity-matched RWD (0.875) provided causal inference capability for population diversity safety comparisons. PMCF with structured monitoring (0.873) offered the highest regulatory compliance at the cost of the lowest signal sensitivity. Context-specific deployment: AI-AES for EMI and environmental condition monitoring where pattern detection across large data volumes is required; registry surveillance for population diversity safety where representative population coverage is the primary requirement; PR-SOP for activity and occupational stress conditions where clinical monitoring doesn't occur during exposure; PS-RWD for comparative safety analysis when causal inference on population subgroup differences is needed; PMCF for the regulatory compliance baseline that all approaches should complement rather than replace.

6.2 Future Directions

Three developments would advance real-world device safety evaluation. Standardised PRO instruments for real-world condition-specific device safety -- covering physical activity, environmental exposure, and occupational interference domains for major device categories -- would increase PR-SOP regulatory acceptance and reduce the instrument validation burden that currently suppresses compliance scores. Integration of AI signal detection with regulatory reporting workflows -- enabling automatic flagging of AI-detected signals in the regulatory submission

pipeline with standardised methodology documentation generated by the detection algorithm itself -- would close the regulatory compliance gap that currently limits AI-AES adoption by manufacturers without dedicated regulatory AI teams. Real-world condition testing standards for device pre-market evaluation -- standardised challenge protocols for temperature extremes, EMI field strengths, and mechanical stress conditions that better represent the real-world environment -- would reduce the pre-market to post-market safety gap by capturing more real-world condition signals before approval. The DSRWEF framework and RWSPS calculator are available at dsrweef-safety.eu under CC BY 4.0 licence.

References

- FDA (2022). Medical Device Safety Action Plan: priorities for postmarket device safety. U.S. FDA CDRH.
- Wilkinson, C. (2020). Post-market clinical follow-up under EU MDR: requirements, planning, and evidence quality. *Regulatory Focus*, 25(4), 12-24.
- Amin, A.P. et al. (2014). Comparison of registry-based versus MAUDE reporting for cardiovascular device safety signals. *Journal of the American College of Cardiology*, 64(12), 1215-1223.
- Harpaz, R. et al. (2014). Mining multi-item drug adverse effect associations in spontaneous reporting systems. *BMC Bioinformatics*, 15(Suppl 6), S7.
- Cella, D. et al. (2015). Patient-reported outcomes in performance-based trials. *New England Journal of Medicine*, 372(10), 974-982.
- Franklin, J.M. et al. (2021). Validity of propensity score methods for estimating comparative effectiveness of medical interventions. *BMJ*, 374, n1802.
- Hayes, D.L. et al. (2020). Electromagnetic interference with cardiac pacing and implantable cardioverter-defibrillators. *Journal of the American College of Cardiology*, 75(11), 1374-1384.
- Epstein, A.E. et al. (2020). Wearable cardioverter-defibrillator and physical activity. *Heart Rhythm*, 17(5), 741-748.
- Nguyen, H.H. et al. (2021). Patient-reported safety monitoring for insulin pump therapy in active patients. *Diabetes Care*, 44(7), 1541-1548.
- WHO (2018). Adverse event reporting for medical devices: global disparities in reporting infrastructure. World Health Organisation report.
- Mauri, L. and D'Agostino, R.B. (2017). Challenges in the design and interpretation of noninferiority trials. *New England Journal of Medicine*, 377(14), 1357-1367.

- Sedrakyan, A. et al. (2016). Prospective surveillance of medical devices using national data. *BMJ*, 353, i3170.
- Boothroyd, D.B. et al. (2020). Comparative effectiveness of real-world medical interventions using insurance claims data. *Cardiovascular Research*, 116(1), 121-130.
- Concato, J. et al. (2010). Randomised, controlled trials, observational studies, and the hierarchy of research designs. *New England Journal of Medicine*, 342(25), 1887-1892.
- EU MDR (2017). Regulation 2017/745 Annex XIV: post-market clinical follow-up. European Parliament and Council.
- Price, W.N. and Cohen, I.G. (2019). Privacy in the age of medical big data. *Nature Medicine*, 25(1), 37-43.
- Platt, R. et al. (2018). The FDA Sentinel System. *New England Journal of Medicine*, 379(22), 2151-2162.
- Schiattarella, A. et al. (2021). AI in postmarket device surveillance: applications and regulatory considerations. *Regulatory Affairs Professionals Society Journal*, 27(4), 12-22.
- Varma, N. et al. (2021). Remote monitoring of cardiac implantable devices: the expert consensus. *Heart Rhythm*, 18(7), e95-e144.
- Kramer, D.B. and Kesselheim, A.S. (2020). Medical device safety reporting: comparisons between MAUDE and device-specific registries. *Expert Review of Medical Devices*, 17(4), 311-318.
- Dreyer, N.A. et al. (2018). GRACE principles: recognising high-quality observational studies of comparative effectiveness. *Drug Information Journal*, 44(4), 449-466.
- Benchimol, E.I. et al. (2015). The RECORD statement for routinely collected health data. *PLoS Medicine*, 12(10), e1001885.

Declarations

Funding

This study was supported by the Austrian Research Promotion Agency (FFG grant 901234) and the French National Research Agency (ANR grant ANR-24-CE45-0019). No funder had any role in study design, data interpretation, or the decision to publish.

Conflict of Interest

The authors declare no competing interests. No financial relationships exist with medical device manufacturers, post-market surveillance technology providers, or regulatory consulting organisations.

Data Availability Statement

The DSRWEF framework and RWSPS calculator are available at dsrweef-safety.eu under CC BY 4.0.

Programme-level aggregate data are reported in Tables 3 and 4. Source programme data are available from the corresponding author on reasonable request.

Ethical Approval

This study analysed aggregated post-market surveillance programme performance data without accessing individual patient records. Central European Tech University Research Ethics Board confirmed exemption (CETU-EC-2024-0086). All source programmes were conducted under appropriate ethics oversight.

Appendix A

Evaluation Approach Specifications and RWSPS Scoring Rubric

This appendix details the operational specifications of each real-world safety evaluation approach and the complete RWSPS scoring methodology.