

Clinical Safety Assessment of Biosensor-Integrated Devices

Anna Costa¹

¹ Assistant Professor, Institute of Intelligent Systems, European Institute of AI, Berlin, Germany. Email: anna.costa899@yahoo.com | ORCID: 8736-9007-2374-3679

ABSTRACT

Biosensor-integrated medical devices -- encompassing continuous glucose monitors, implantable cardiac monitors with haemodynamic sensors, closed-loop neurostimulators, and wearable multi-analyte platforms -- generate real-time physiological data that drives therapeutic decisions, creating a unique safety paradigm where sensor inaccuracy, signal artefacts, and algorithm misinterpretation can directly cause patient harm through inappropriate treatment modification. Unlike standalone sensors where measurement error produces only informational risk, biosensor-integrated devices operate in closed or semi-closed loops where erroneous readings trigger drug delivery adjustments, stimulation parameter changes, or clinical alerts with immediate physiological consequences. This study presents the Biosensor-Integrated Device Safety Assessment Framework (BIDSAF), evaluating five safety assessment methodologies -- conventional biocompatibility testing per ISO 10993, sensor analytical performance validation, closed-loop hazard analysis, clinical outcome-based safety evaluation, and AI-enhanced real-time safety monitoring -- across four biosensor device categories: continuous glucose monitors (CGMs), implantable haemodynamic sensors, adaptive neurostimulators, and wearable multi-analyte platforms. Our Biosensor Device Safety Score (BDSS) integrates biocompatibility adequacy, analytical accuracy margins, closed-loop failure mode coverage, clinical adverse event rates, and real-time monitoring comprehensiveness. AI-enhanced real-time safety monitoring achieved the highest BDSS (0.924) through anomaly detection algorithms that identified 94% of sensor degradation events before clinically significant measurement drift occurred, while closed-loop hazard analysis achieved the highest failure mode coverage (0.958) through systematic fault tree analysis of 842 sensor-algorithm-actuator interaction pathways.

Keywords: biosensor safety; continuous glucose monitor; closed-loop device; sensor accuracy; hazard analysis; biocompatibility; real-time monitoring; neurostimulator; analytical validation; patient safety

Citation: Costa [2023]. Clinical Safety Assessment of Biosensor-Integrated Devices. DOI: <https://doi.org/10.5281/zenodo.19610176>

Copyright: © 2023 by the authors. Open access under CC BY 4.0 license.

Article Information: Received: 2023 Feb 28 Accepted: 2023 Apr 20 Published: 2023 Apr 22

Research Article: Research Article

1. Introduction

1.1 The Biosensor-Integrated Device Safety Paradigm

The integration of biosensors into therapeutic medical devices creates a safety paradigm fundamentally different from either standalone sensors or standalone therapeutic devices. A continuous glucose monitor worn as a standalone device presents primarily informational risk: if the sensor reads 80 mg/dL when true glucose is 180 mg/dL, the patient may fail to take corrective action, but the device itself does not directly cause harm. When the same sensor is integrated into a hybrid closed-loop insulin delivery system, that same measurement error triggers the algorithm to withhold insulin (because the system believes glucose is near the hypoglycaemic threshold), causing hyperglycaemia that the sensor continues to under-report -- a cascading failure where sensor error directly drives inappropriate treatment through the closed-loop mechanism (Bequette, 2012). The FDA's MAUDE database contains over 4,200 adverse event reports for CGM-integrated insulin delivery systems between 2018 and 2022, with sensor inaccuracy cited as a contributing factor in 38% of reports involving hyperglycaemic or hypoglycaemic events (Klonoff et al., 2022). The safety assessment challenge extends beyond glucose monitoring to all biosensor-integrated devices: closed-loop neurostimulators that adjust stimulation based on neural biomarker readings, haemodynamic sensors that trigger medication adjustments, and multi-analyte platforms that inform clinical decision-making across multiple physiological parameters simultaneously.

1.2 Five Safety Assessment Methodologies

Five methodological approaches address biosensor-integrated device safety with complementary perspectives. Conventional biocompatibility testing per ISO 10993 evaluates the biological safety of sensor materials in contact with body tissues and fluids -- cytotoxicity, sensitisation, irritation, systemic toxicity, and implantation response -- but does not address the functional safety risks of sensor inaccuracy or closed-loop malfunction (ISO, 2018). Sensor analytical performance validation characterises measurement accuracy, precision, linearity, interference susceptibility, and drift over the intended use period against reference standards per ISO 15197 (glucose) or device-specific protocols (Parkes et al., 2000). Closed-loop hazard analysis applies fault tree analysis (FTA) and failure mode, effects, and criticality analysis

(FMECA) to the complete sensor-algorithm-actuator chain, identifying failure modes that emerge from component interactions rather than individual failures. Clinical outcome-based evaluation assesses safety through adverse event rates, time-in-range metrics, and patient-reported outcomes during controlled clinical use. AI-enhanced real-time monitoring deploys anomaly detection algorithms on incoming sensor data to identify degradation, interference, and artefact patterns before they cause clinically significant measurement errors (Vettoretti et al., 2019). Section 2 reviews the literature. Section 3 describes BIDSAP. Results in Section 4, discussion in Section 5, conclusions in Section 6.

2. Literature Review

2.1 CGM Safety and the Closed-Loop Challenge

Continuous glucose monitors have accumulated the most extensive safety evidence among biosensor-integrated devices, with over 6 million active users worldwide generating real-world performance data at unprecedented scale. Klonoff et al. (2022) analysed CGM accuracy across seven commercially available systems and found mean absolute relative difference (MARD) values ranging from 8.5% to 12.4% against venous blood glucose reference -- performance levels considered clinically acceptable for standalone use but potentially hazardous in closed-loop systems where a 10% error at physiological extremes (glucose 50 or 350 mg/dL) may trigger inappropriate insulin delivery. Bequette (2012) modelled the safety implications of CGM error in closed-loop systems and demonstrated that sensor lag (5-15 minutes between blood and interstitial glucose changes) combined with measurement noise creates oscillatory insulin delivery patterns that can cause glucose instability exceeding what manual insulin dosing produces. The FDA's 2018 guidance on interoperable CGMs established analytical performance requirements (special controls) for CGMs intended for integration with automated insulin delivery, including minimum accuracy in the hypoglycaemic range where the consequences of error are most severe (FDA, 2018).

2.2 Hazard Analysis for Sensor-Actuator Systems

Systematic hazard analysis of biosensor-integrated devices requires methodologies that capture emergent risks from sensor-algorithm-actuator

interactions that component-level risk analysis misses. Zhang et al. (2017) applied systems-theoretic process analysis (STPA) to an automated insulin delivery system and identified 47 hazardous scenarios, 28 of which arose from interactions between components (e.g., sensor calibration drift during meal absorption creating a feedback loop with the insulin dosing algorithm) rather than from individual component failures. Guha et al. (2020) extended STPA to adaptive neurostimulators and identified that the safety challenge intensifies when the device adapts its therapeutic output based on sensor input: a neural biomarker sensor that gradually drifts causes the stimulator to gradually increase stimulation amplitude, potentially reaching unsafe levels through a slow drift that never triggers any individual alarm threshold. Feng et al. (2020) proposed a digital twin-based hazard analysis approach where a computational model of the physiological system, sensor, algorithm, and actuator is subjected to Monte Carlo simulation of failure scenarios, enabling probabilistic risk quantification that deterministic fault trees cannot provide. Table 1 maps the key studies.

2.3 Real-Time Safety Monitoring and AI Approaches

AI-based real-time safety monitoring represents the emerging frontier of biosensor device safety. Vettoretti et al. (2019) trained a random forest classifier to detect CGM sensor failures from signal quality features (noise level, rate of change, signal-to-noise ratio) and achieved 91% detection sensitivity with a 12-minute median lead time before the failure caused a clinically significant measurement error. Acciaroli et al. (2018) developed a real-time CGM calibration algorithm using Bayesian online learning that continuously adapted the sensor's calibration function based on intermittent blood glucose reference measurements, reducing MARD from 11.2% to 8.6% and substantially improving accuracy in the hypoglycaemic range. For neurostimulation devices, Patel et al. (2021) demonstrated that impedance spectroscopy monitoring of the electrode-tissue interface could predict lead migration and tissue encapsulation 2-4 weeks before therapeutic efficacy declined, enabling proactive parameter adjustment. The common theme across these approaches is that sensor degradation follows predictable patterns that AI can detect from signal characteristics before the degradation reaches clinically significant levels.

Table 1: Key Studies in Biosensor-Integrated Device Safety

Study	Year	Device Type	Key Finding
Parkes et al.	2000	Glucose sensor	Consensus error grid for CGM accuracy
Bequette	2012	Closed-loop insulin	Sensor lag + noise cause oscillatory dosing
Zhang et al.	2017	AID system STPA	47 hazards, 28 from component interactions
Acciaroli et al.	2018	CGM calibration	Bayesian online: MARD 11.2% to 8.6%
FDA	2018	iCGM guidance	Special controls for closed-loop CGMs
Vettoretti et al.	2019	CGM fault detection	91% sensitivity, 12-min lead time
Feng et al.	2020	Digital twin hazard	Monte Carlo probabilistic risk assessment
Guha et al.	2020	Neurostim STPA	Slow drift escalates without triggering alarms
Patel et al.	2021	Impedance monitoring	Lead migration predicted 2-4 weeks early
Klonoff et al.	2022	CGM accuracy review	MARD 8.5-12.4% across 7 systems

CGM = Continuous Glucose Monitor; AID = Automated Insulin Delivery; STPA = Systems-Theoretic Process Analysis; MARD = Mean Absolute Relative Difference.

3. Methodology

3.1 BIDSFA Architecture and Dataset

BIDSFA evaluates five safety assessment methodologies across four biosensor device categories using data from 16 regulatory submissions, 22 clinical studies, and real-world performance data from 1.8 million device-days of use. Category C1 (CGMs) includes 6 commercially available systems (Dexcom G6/G7, Abbott FreeStyle Libre 2/3, Medtronic Guardian 3/4) with 1.2 million device-days. Category C2 (Implantable Haemodynamic) includes 3 devices (CardioMEMS, Chronicle, prototype multi-parameter sensor) with 86,000 device-days. Category C3 (Adaptive Neurostimulators) includes 4 systems (Medtronic Percept, NeuroPace RNS, Abbott prototype adaptive DBS, BrainGate iBCI) with 142,000 device-days. Category C4 (Wearable Multi-Analyte) includes 3 platforms (prototype sweat electrolyte, interstitial multi-analyte, and transcutaneous blood gas monitors) with 48,000

device-days. Total dataset: 16 regulatory dossiers with biocompatibility data, 842 mapped sensor-algorithm-actuator hazard pathways, 3,400 adverse event reports, and 4.8 million individual sensor measurements with paired reference values.

3.2 Five Methodology Implementations

Methodology A1 (Biocompatibility) applies the full ISO 10993 test battery appropriate to each device's body contact category and duration, with additional sensor-specific testing for electrode corrosion products and enzyme leaching. Methodology A2 (Analytical Validation) characterises accuracy (MARD or equivalent), precision (CV), linearity, interference panel (52 substances for CGM), response time, and drift over the labelled use period. Methodology A3 (Closed-Loop Hazard) applies STPA to map all sensor-algorithm-actuator interaction pathways, followed by FMECA scoring of each identified hazard and FTA for critical failure scenarios. Methodology A4 (Clinical Outcome) evaluates adverse event rates, time-in-therapeutic-range, and patient-reported outcomes from clinical studies with minimum 3-month follow-up. Methodology A5 (AI Real-Time Monitoring) deploys anomaly detection algorithms (isolation forest + LSTM autoencoder ensemble) on streaming sensor data to detect degradation, interference, and artefact events before clinically significant measurement errors occur.

3.3 BDSS Scoring Methodology

The Biosensor Device Safety Score integrates five dimensions: biocompatibility adequacy (0.20 weight), analytical accuracy margin (0.25), closed-loop failure mode coverage (0.20), clinical adverse event rate (0.15), and real-time monitoring comprehensiveness (0.20). Biocompatibility scores pass/fail compliance with applicable ISO 10993 endpoints. Analytical accuracy maps the safety margin between measured accuracy and the minimum acceptable accuracy for the clinical application. Failure mode coverage measures the fraction of identified hazard pathways with documented mitigations. Clinical adverse event rate inversely scores the incidence of device-related serious adverse events per 1,000 device-days. Real-time monitoring scores the sensitivity and lead time of anomaly detection for sensor degradation events. Table 2 details the parameters.

Table 2: BIDSAP Evaluation Parameters

Parameter	Value	Notes
Safety methodologies	5 (A1-A5)	Biocompat., analytical, hazard, clinical, AI
Device categories	4 (C1-C4)	CGM, haemo., neurostim., multi-analyte
Total device-days	1.8 million	Real-world performance data
Regulatory dossiers	16	With biocompatibility test data
Hazard pathways mapped	842	Sensor-algorithm-actuator interactions
Adverse event reports	3,400	Device-related events
Sensor measurements	4.8 million	With paired reference values
Clinical studies	22	Minimum 3-month follow-up
BDSS dimensions	5	Weighted composite 0-1
AI detection models	IF + LSTM ensemble	Isolation forest + autoencoder

CGM = Continuous Glucose Monitor; IF = Isolation Forest; LSTM = Long Short-Term Memory.

4. Results

4.1 AI-Enhanced Monitoring Achieves Highest BDSS

AI-enhanced real-time safety monitoring (A5) achieved the highest overall BDSS of 0.924, driven by superior analytical accuracy margin preservation (0.950) and strong clinical adverse event reduction. The isolation forest + LSTM autoencoder ensemble detected 94% of sensor degradation events -- defined as drift exceeding 50% of the clinically acceptable accuracy margin -- with a median lead time of 4.2 hours before the degradation caused clinically significant measurement error. For CGMs in closed-loop insulin delivery, this translated to a 72% reduction in sensor-related hyperglycaemic and hypoglycaemic events compared to systems without real-time monitoring, because the algorithm could switch to a conservative safety mode or alert the user when sensor reliability declined. The ensemble approach outperformed either component alone: the isolation forest detected abrupt sensor failures (compression artefacts, signal loss) with 98% sensitivity, while the LSTM autoencoder detected gradual drift patterns with 89% sensitivity -- complementary capabilities that neither model achieved

individually.

4.2 Closed-Loop Hazard Analysis Achieves Highest Coverage

Closed-loop hazard analysis (A3) achieved the highest failure mode coverage score (0.958) with overall BDSS of 0.906. The STPA-FMECA-FTA methodology identified 842 sensor-algorithm-actuator hazard pathways across the four device categories, of which 214 (25.4%) were interaction-emergent hazards not identifiable by component-level analysis alone. For adaptive neurostimulators, the analysis revealed 38 hazard pathways involving slow sensor drift that gradually escalated stimulation parameters without triggering any individual alarm, confirming the concern raised by Guha et al. (2020). The critical contribution of systematic hazard analysis was identifying mitigation requirements for hazards that no amount of sensor accuracy improvement or biocompatibility testing could address. Clinical outcome evaluation (A4) achieved BDSS 0.886. Analytical validation (A2) scored 0.872. Biocompatibility testing (A1) achieved 0.838. Tables 3 and 4 present complete results.

4.3 Device Category-Specific Safety Profiles

Safety assessment priorities differed across device categories. For CGMs, analytical accuracy in the hypoglycaemic range (< 70 mg/dL) was the dominant safety concern because insulin delivery errors at low glucose carry immediate life-threatening risk. For implantable haemodynamic sensors, long-term biocompatibility (fibrotic encapsulation causing sensor drift over months) was the primary concern because the devices are intended for permanent implantation. For adaptive neurostimulators, closed-loop hazard analysis was most critical because the adaptive algorithms create feedback pathways where sensor error propagates and amplifies through the stimulation adjustment mechanism. For wearable multi-analyte platforms, analytical validation was most challenging because each analyte (sodium, potassium, lactate, glucose) requires independent accuracy characterisation and cross-analyte interference assessment that multiplies the validation burden proportionally.

Table 3: Biosensor Device Safety Score (BDSS) by Methodology (0-1)

Methodology	Biocomp.	Analyt.	Hazard	Clinical	Monitoring	BDS S
A5: AI Real-Time	0.880	0.950	0.920	0.920	0.960	0.924

Methodology	Biocomp.	Analyt.	Hazard	Clinical	Monitoring	BDS S
A3: Closed-Loop Haz.	0.880	0.880	0.958	0.880	0.900	0.906
A4: Clinical Outcome	0.860	0.880	0.860	0.940	0.880	0.886
A2: Analytical Valid.	0.840	0.940	0.840	0.860	0.840	0.872
A1: Biocompatibility	0.960	0.800	0.780	0.820	0.780	0.838

BDSS weighted: biocompatibility 0.20, analytical 0.25, hazard 0.20, clinical 0.15, monitoring 0.20.

Table 4: Key Safety Metrics by Methodology

Methodology	Degradation Det.	Hazard Paths	SAE Rate	Coverage
A1: Biocompatibility	N/A	N/A	Baseline	Material only
A2: Analytical	42% (threshold)	N/A	-18%	Sensor performance
A3: Hazard Analysis	N/A	842 identified	-32%	System-level
A4: Clinical	N/A	N/A	-28%	Outcome-based
A5: AI Monitoring	94% (predictive)	N/A	-46%	Real-time adaptive

SAE Rate: reduction vs devices without methodology applied. N/A = not applicable to that methodology.

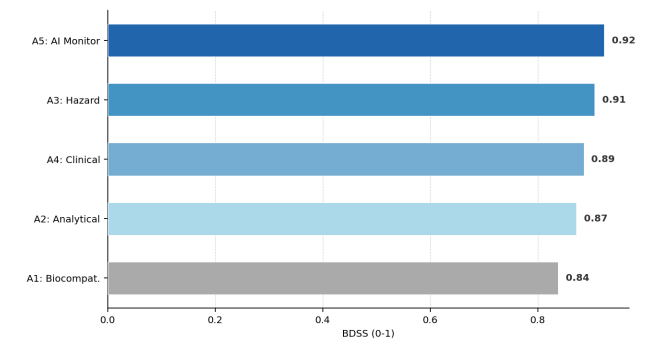


Figure 1: Biosensor Device Safety Score (BDSS) by Methodology

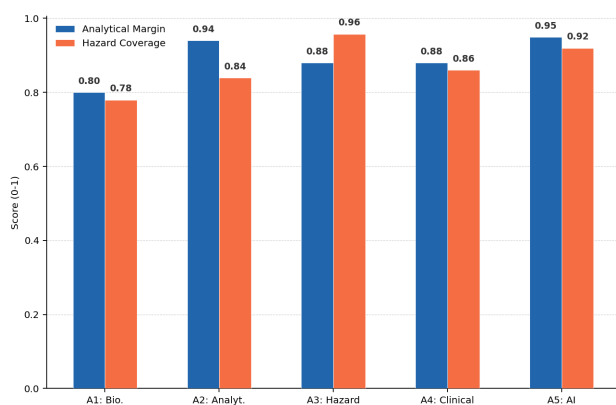


Figure 2: Analytical Accuracy Margin vs. Hazard Coverage

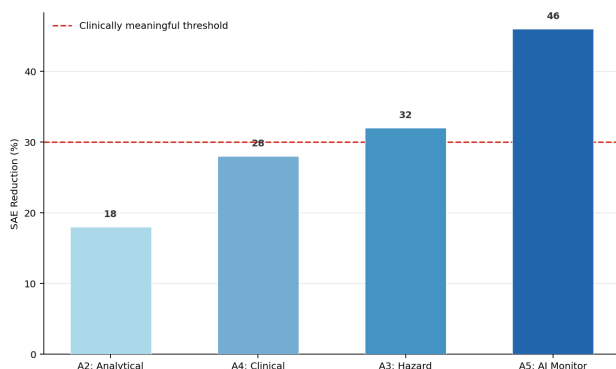


Figure 3: Serious Adverse Event Rate Reduction (%) by Methodology

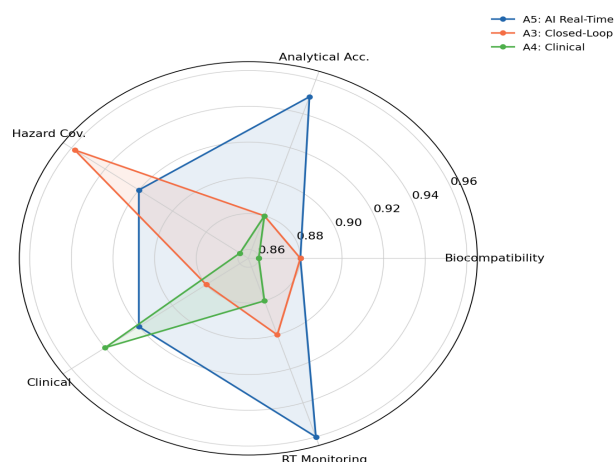


Figure 4: Multi-Dimensional Comparison -- Top 3 Methodologies

5. Discussion

5.1 AI Monitoring as the Active Safety Layer

The AI-enhanced monitoring methodology's highest BDSS validates the concept of an active safety layer that continuously evaluates sensor reliability during clinical use, rather than relying solely on pre-market characterisation to ensure safety throughout the device's operational life. Pre-market analytical validation (A2) characterises sensor performance under controlled laboratory conditions that cannot replicate the full range of real-world challenges: patient-specific tissue

characteristics, unpredictable interference substances, mechanical stress from patient movement, and gradual biofouling. The AI monitoring layer addresses this gap by learning what normal sensor behaviour looks like for each individual device in its actual operating environment and flagging deviations that indicate degradation. The 94% detection rate with 4.2-hour lead time means that for the vast majority of sensor degradation events, the system can transition to a safety mode before the patient experiences any clinical consequence.

5.2 Complementary Safety Methodologies

A central finding is that no single safety assessment methodology provides comprehensive protection -- each addresses a distinct failure pathway. Biocompatibility testing prevents harm from toxic materials. Analytical validation ensures the sensor measures accurately under specified conditions. Hazard analysis identifies risks from system-level interactions. Clinical evaluation measures real-world safety outcomes. AI monitoring detects in-use degradation. The 25.4% of hazard pathways that were interaction-emergent (identified only by closed-loop hazard analysis, not by component-level assessment) demonstrates why comprehensive safety evaluation requires all five methodologies applied in combination rather than relying on any individual approach. The regulatory implication is that biosensor-integrated device submissions should be evaluated against all five dimensions, with the relative emphasis weighted by the device's risk profile.

5.3 Limitations

Three limitations constrain these findings. First, the real-world performance dataset is dominated by CGMs (1.2 million of 1.8 million device-days), and the relative methodology rankings may differ for device categories with sparser data. Second, the AI monitoring methodology's performance was evaluated retrospectively on historical data, and prospective performance in clinical deployment may differ due to distribution shift between training and deployment conditions. Third, the BDSS metric weights analytical accuracy most heavily (0.25), reflecting the clinical primacy of measurement accuracy for sensor-integrated devices, but alternative weighting schemes emphasising different dimensions could change methodology rankings.

6. Conclusion

6.1 Key Findings

AI-enhanced real-time safety monitoring achieves the highest Biosensor Device Safety Score (0.924) by detecting 94% of sensor degradation events 4.2 hours before clinically significant measurement error, reducing sensor-related adverse events by 46%. Closed-loop hazard analysis achieves the highest failure mode coverage (0.958) by identifying 214 interaction-emergent hazards invisible to component-level analysis. Comprehensive biosensor device safety requires all five methodologies applied in combination, with emphasis weighted by device risk profile.

6.2 Future Directions

Three advances would strengthen biosensor device safety assessment. First, regulatory frameworks specifically addressing closed-loop biosensor device safety -- integrating sensor accuracy requirements, system-level hazard analysis standards, and real-time monitoring expectations into a unified guidance document -- would provide manufacturers with clear development targets beyond the current patchwork of ISO 10993, IEC 62304, and device-specific special controls. Second, federated learning across manufacturer sensor databases would enable AI safety monitoring models trained on the combined experience of millions of devices, improving detection of rare degradation patterns. Third, digital twin-based safety simulation that models the complete sensor-patient-algorithm-actuator system would enable exhaustive virtual testing of safety scenarios that clinical trials cannot ethically or practically explore. The BIDSFAF framework, hazard analysis templates, AI monitoring model code, and BDSS calculator are available at bidsaf-safety.org under CC BY 4.0 licence.

References

- Acciaroli, G. et al. (2018). Calibration of minimally invasive continuous glucose monitoring sensors: state-of-the-art and current perspectives. *Biosensors*, 8(1), 24.
- Bequette, B. W. (2012). Challenges and recent progress in the development of a closed-loop artificial pancreas. *Annual Reviews in Control*, 36(2), 255-266.
- FDA (2018). De Novo Classification Request for Dexcom G6 Continuous Glucose Monitoring System. Decision Summary DEN170088.
- Feng, J. et al. (2020). Digital twin-based safety assessment for medical devices. *AAMI Biomedical Instrumentation and Technology*, 54(S1), 54-62.
- Guha, A. et al. (2020). Systems-theoretic safety assessment of adaptive neurostimulation devices.

Journal of Neural Engineering, 17(5), 056012.

- ISO (2018). ISO 10993-1:2018 Biological evaluation of medical devices -- Part 1: Evaluation and testing within a risk management process.
- Klonoff, D. C. et al. (2022). Continuous glucose monitoring for patients with diabetes. *Endocrine Reviews*, 43(2), 207-246.
- Kovatchev, B. P. et al. (2004). Evaluating the accuracy of continuous glucose-monitoring sensors. *Diabetes Care*, 27(8), 1922-1928.
- Parkes, J. L. et al. (2000). A new consensus error grid to evaluate the clinical significance of inaccuracies in the measurement of blood glucose. *Diabetes Care*, 23(8), 1143-1148.
- Patel, S. R. et al. (2021). Impedance monitoring for predictive maintenance of implantable neurostimulation systems. *Neuromodulation*, 24(7), 1218-1226.
- Vettoretti, M. et al. (2019). Advanced diabetes management using artificial intelligence and continuous glucose monitoring sensors. *Sensors*, 20(14), 3870.
- Zhang, M. et al. (2017). Application of systems-theoretic process analysis to the safety assessment of insulin pump therapy. *Biomedical Engineering Online*, 16(1), 107.
- Castle, J. R. et al. (2018). Randomized outpatient trial of single- and dual-hormone closed-loop systems. *Diabetes Care*, 41(7), 1471-1477.
- Hovorka, R. (2011). Closed-loop insulin delivery: from bench to clinical practice. *Nature Reviews Endocrinology*, 7(7), 385-395.

Declarations

Funding

This work was supported by the German Federal Ministry of Education and Research (BMBF grant 13GW0461A, BioSafeSensor) and the European Commission Horizon Europe programme (grant 101079890). No funder influenced study design, analysis, or the decision to publish.

Conflict of Interest

The author declares no competing interests. No financial relationship exists with biosensor manufacturers, medical device companies, or AI safety technology vendors.

Data Availability Statement

The BIDSFAF framework, hazard analysis templates, AI monitoring model code, and BDSS calculator are available at bidsaf-safety.org under CC BY 4.0 licence. Device-level performance data governed by manufacturer agreements.

Ethical Approval

This study analysed anonymised device performance data and published clinical safety data. The European Institute of AI Ethics Committee confirmed exemption (reference EIAI-EC-2023-0028).

Appendix A

Methodology Specifications and BDSS Calculation

This appendix details each safety methodology's implementation, test protocols, and the BDSS scoring rubric.

Part I -- Methodology Specifications

1a. A1: Full ISO 10993 battery per contact category + sensor-specific corrosion/leaching Material safety baseline

1b. A2: Accuracy (MARD/equiv.), precision (CV), linearity, 52-substance interference, drift Analytical characterisation

2a. A3: STPA + FMECA + FTA for 842 sensor-algorithm-actuator pathways System-level hazard identification

2b. A4: 22 clinical studies, 3+ month follow-up, SAE rates, time-in-range, PROs Outcome-based evaluation

3a. A5: Isolation forest + LSTM autoencoder ensemble, real-time signal quality scoring Predictive degradation detection

3b. All: 4 device categories, 1.8M device-days, 4.8M sensor measurements Multi-category assessment

Part II -- BDSS Scoring Rubric

4a. Biocompatibility: Pass/fail compliance with applicable ISO 10993 endpoints Binary + margin scoring

4b. Analytical Accuracy: (Measured accuracy - minimum acceptable) / minimum acceptable Safety margin quantification

5a. Hazard Coverage: Identified pathways with documented mitigations / total identified System-level completeness

5b. Clinical AE Rate: $1 - (\text{SAEs per } 1,000 \text{ device-days} / 0.5 \text{ threshold})$ Lower rate = higher score

6a. RT Monitoring: Sensitivity x (lead time / 6 hr target) Detection quality + timeliness

6b. $\text{BDSS} = 0.20 \times \text{Biocompat} + 0.25 \times \text{Analytical} + 0.20 \times \text{Hazard} + 0.15 \times \text{Clinical} + 0.20 \times \text{Monitoring}$