AI in PMS: Is this even compliant?

Tibor Zechmeister, B.Sc. MSc EMBA

October 17th, 2024

Personal Introduction

- 12 years of experience in regulatory affairs for medical devices
- Founded and co-founded medical device companies (Neurodegenerative Diseases)
- Acted as manufacturer, PRRC, Head of regulatory
- External Auditor at Notified Body 1304
- Head of Regulatory and Quality at Flinn.ai

Tibor Zechmeister
B.Sc. M.Sc. EMBA

Flinn Introduction

- Software as a Service provider for regulatory compliance

- Main focus on PMS:
  - Vigilance Database Monitoring
  - Literature Search
  - Complaint Handling
  - Risk Management
  - Regulatory Monitoring

- Automation approaches, but also integration of AI features for users

The Overall Goal of PMS

- Compliance with Regulations
- Collecting Data about Device Effectiveness

- At the end of the day: Patient Safety

The Current Reality with PMS

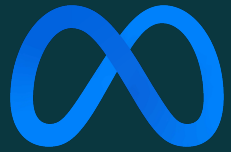
- Manual Data Processing
- Extensive Literature Searches
- Limited Resources due to high workload

Examples of Intensive Activities

- Screening of scientific literature and other sources of clinical data
- Post-market studies
- Monitoring vigilance databases
- Survey from health care professional / user
- Complaint handling
- Review of case reports which may reveal misuse or off-label use

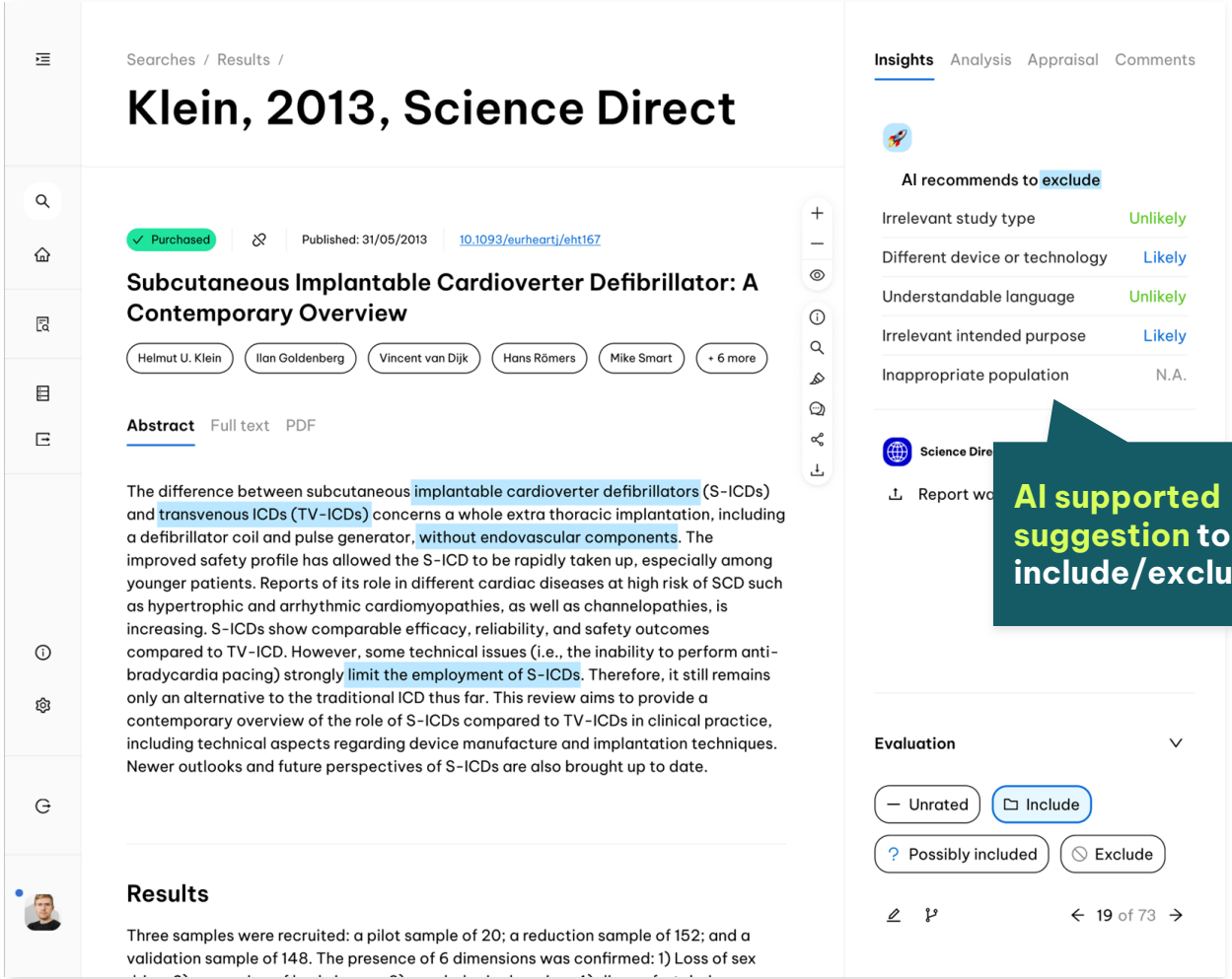
Can AI help?

- Is AI in PMS a long-term solution or just a hype?
- What parts of PMS could be supported?
- What are current solutions and future approaches?



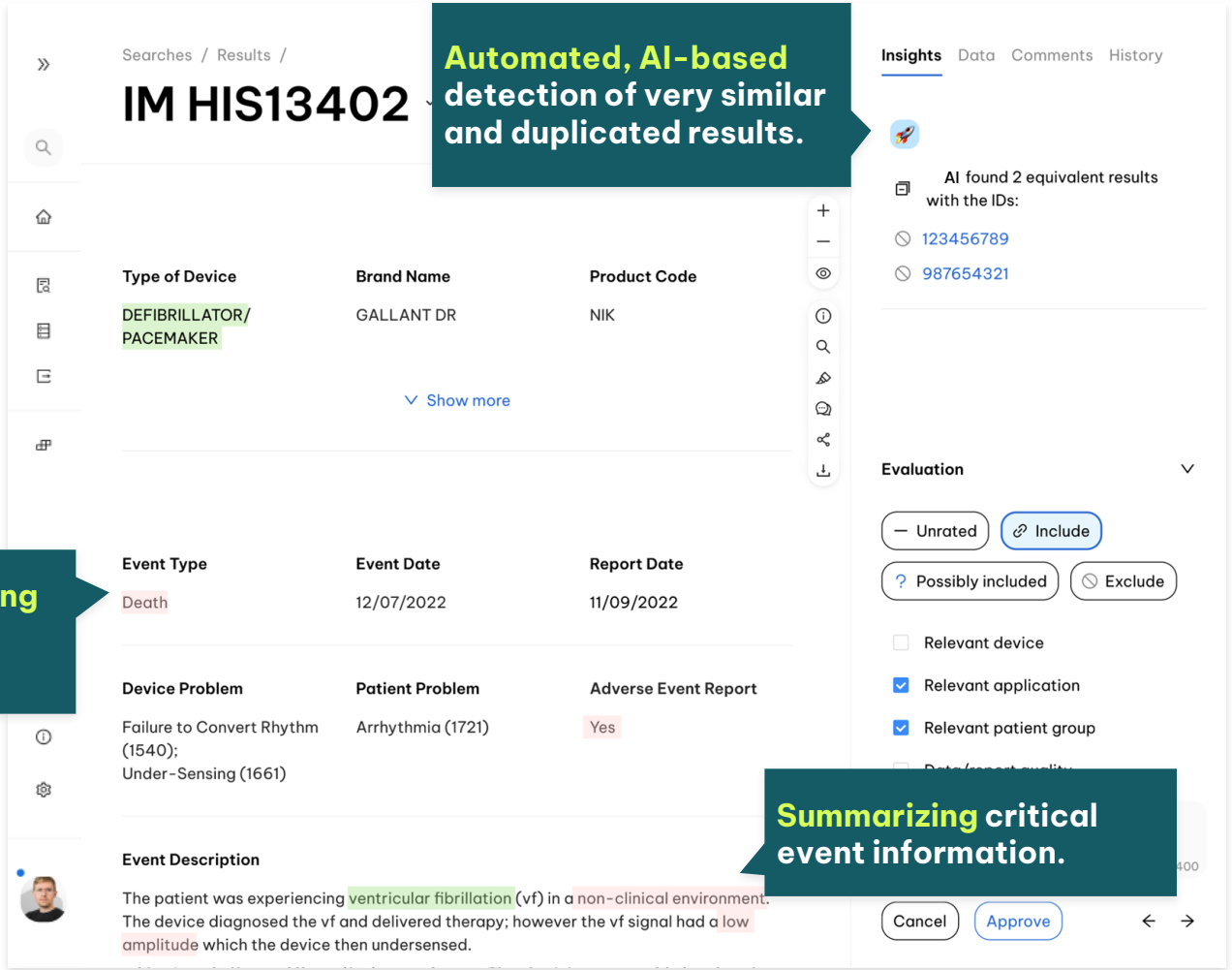
AI in PMS: Is this even compliant?

Example of Literature Search



AI in PMS: Is this even compliant?

Example of Survey Evaluation



Future Possible Approaches

The Main Challenges with AI in PMS

Potential Auditor Questions:

1. Will the system hallucinate?
2. How was the system validated?
3. What data sets were used for training?

The Main Challenges with Validation

- Every Use Case needs a different approach of validation
- Even Large Language Models struggle with sometimes simple tasks (Identify death in reports)
- This type of validation needs much more than regulatory expertise: Data Scientists and Prompt Engineers are needed

Examples of Validation Approaches

- **1. Performance-Based Validation**
- **Benchmarking**: Using established datasets and benchmarks (e.g., GLUE, SuperGLUE, SQuAD) to evaluate the model's performance.
- **Cross-validation**: Splitting the data into training and testing sets multiple times to ensure consistent performance.
- **A/B Testing**: Comparing two versions of the model to see which performs better in real-world scenarios.
- **2. Robustness Testing**
- **Adversarial Testing**: Assessing the model's resilience against adversarial inputs designed to trick or mislead it.
- **Stress Testing**: Evaluating the model under extreme conditions or unusual inputs to observe its behavior.
- **3. Fairness and Bias Evaluation**
- **Bias Detection**: Identifying and quantifying biases in the model's predictions across different demographic groups.
- **Fairness Audits**: Conducting audits to ensure the model treats all user groups fairly.
- **4. Explainability and Interpretability**
- **Model Explainability**: Using techniques like SHAP (SHapley Additive exPlanations) or LIME (Local Interpretable Model-agnostic Explanations) to make the model's predictions understandable.
- **Transparency Reports**: Documenting the model's decision-making process and potential limitations.
- **5. Ethical and Responsible AI Evaluation**
- **Ethical AI Guidelines**: Adhering to established guidelines and principles for ethical AI development and deployment.
- **Impact Assessments**: Evaluating the societal and ethical impacts of deploying the model.
- **6. Security Testing**
- **Vulnerability Analysis**: Identifying potential security vulnerabilities that could be exploited.
- **Penetration Testing**: Simulating attacks to test the model's security measures.
- **7. Usability Testing**
- **User Feedback**: Collecting feedback from end-users to evaluate the model's effectiveness and usability.
- **Human-in-the-Loop**: Involving human oversight in the model's decision-making process to ensure accuracy and reliability.
- **8. Scalability and Efficiency Testing**
- **Scalability Testing**: Assessing the model's performance when scaled up to handle large volumes of data or concurrent users.

Just to name a few..

Regulatory Requirements

- MDR
  - The MDR does not address AI at all
- EU AI Act
  - Provides general information and validation data quality and procedures, but no concrete approaches
- Standards
  - Multiple guidance documents and standards are available, such as BS 30440
  - Still requirements are on a general level

Example 1: Detecting Keywords in Reports

- Ground Truth testing
- System is tasked to identify certain keywords
- Human expert checks if decisions are correct
- How many checks are done? 100% would not be feasible
- Proper proportion needs to be identified based on number of results

Example 2: Categorizing Events

- Complaint Handling categorization
- Here, there are issues with the ground truth, especially with multiple experts
- Example of categorizing events or reports with IMDRF coding

Summary

- PMS is a promising field of implementing AI
- Current approaches and solutions are already available or being developed
- Manufacturers still need to be aware about the multiple challenges regarding reliability and validation
- Manufacturers and service providers need to have a multidisciplinary team of regulatory experts and data scientists

Would you use AI in your PMS activities?