

## Appendix: Content Summaries of Selected Best Papers for the IMIA Yearbook 2023 Section "Clinical Information Systems"<sup>1</sup>

Bialke M, Geidel L, Hampf C, Blumentritt A, Penndorf P, Schuldt R, Moser FM, Lang S, Werner P, Stäubert S, Hund H, Albashiti F, Gührer J, Prokosch HU, Bahls T, Hoffmann W

**A FHIR has been lit on gICS: facilitating the standardised exchange of informed consent in a large network of university medicine**

BMC Med Inform Decis Mak 2022 Dec 19;22(1):335. doi: 10.1186/s12911-022-02081-4

### Overview:

The article discusses the technical implementation of patient consent management in the context of the NUM-CODEX project, which aims to enable cross-site data exchange via FHIR in the area of consent. The authors describe the use of globally unique object identifiers (OID) and semantic statements to develop a common representation of the MII Broad Consent, which was extended with NUM-specific extensions. They also highlight the challenges of implementing complex FHIR consent profiles and the need for extensive practical experience and quality control to ensure syntactically and semantically correct implementation.

### Detailed Description:

#### Goal

The goal of the text is to describe the implementation of a FHIR-compliant and interoperable nationwide exchange of consent information using gICS and TTP-FHIR Gateway.

#### Methods

The solution covers requirements identified in NUM-CODEX setting, which is part of a network of university medicines (NUM) to support COVID-19 and pandemic research at national level.

All 34 participating university hospitals work upon harmonized infrastructural as well as legal basis for their data protection compliant collection and transfer.

Informed consent from patients is required for processing their health data at NUM sites, transfer to CODEX, use & access procedure based on GDPR Art. 6(1) lit. a).

#### Findings

A reliable consent management system was successfully implemented by University Medicine Greifswald using MII broad consent in mid2020.

Current developments in the FHIR community must be considered when implementing cross-site data exchange via FHIR consents.

Steps were taken towards conception/implementation standardized solutions for provisioning consents through gICS with TTP-FHIT gateway serving as an intermediary between external infrastructure components.

#### Implications

Technical prerequisites have been achieved with help from gICS/TTP-FHIT Gateway enabling FHIT compliant provision info about patient's informed consents across multiple sites while maintaining compliance w/GDPR regulations.

Customised templates simplify assurance technical interoperability among all NUM/ MII sites.

Guardiolle V, Bazoge A, Morin E, Daille B, Toublant D, Bouzillé G, Merel Y, Pierre-Jean M, Filiot A, Cuggia M, Wargny M, Lamer A, Gourraud PA

**Linking Biomedical Data Warehouse Records With the National Mortality Database in France: Large-scale Matching Algorithm**

JMIR Med Inform. 2022 Nov 1;10(11):e36711. doi: 10.2196/36711

### Overview:

The paper describes a large-scale matching algorithm that links biomedical data warehouse (BDW) records with the French National Mortality Database (FNMD) to determine vital status after discharge, which is crucial for medical research. The algorithm uses advanced data cleaning and knowledge of the naming system, along with the Damerau-Levenshtein distance (DLD), to overcome challenges such as absence of unique common identifiers, name changes, and clerical errors. The algorithm's performance was evaluated using BDW

data from three university hospitals, and the results showed that the DLD-based algorithm outperformed the direct algorithm in terms of sensitivity/recall by 11%.

### Detailed Description:

#### Goal

The goal of the text is to present a study that developed and evaluated two algorithms for linking patient records between local biomedical data warehouses (BDWs) and the French National Mortality Database (FNMD).

#### Methods

The study utilized record linkage techniques, advanced data cleaning, knowledge of naming systems, Damerau-Levenshtein distance calculations, and blocking techniques.

Two algorithms were developed: a direct-matching algorithm and a deterministic matching algorithm based on DLD.

#### Findings

Matching large-scale BDW records with FNMD is challenging due to absence of unique common identifiers between databases.

The use of an advanced deterministic matching algorithm such as the DLD-based algorithm showed higher sensitivity/recall than direct matching algorithms.

Specificity was  $\geq 98\%$ , reducing risk differential biases between groups.

The execution time per patient decreased as total number increased but blocking technique reduced required comparisons by at least 40k times.

#### Implications

The findings have implications for medical research using open-source external data sources like FNMD in improving usage value BDWs.

The proposed method can be used routinely update vital status in BDW records from FNMD on large scale

Deterministic methods are more efficient when dealing with datasets without unique common identifiers

Poelzl G, Egelseer-Bruendl T, Pfeifer B, Modre-Osrian R, Welte S, Fetz B, Krestan S, Haselwanter B, Zaruba MM, Doerfler J, Rissbacher C, Ammenwerth E, Bauer A

**Feasibility and effectiveness of a multidimensional post-discharge disease management programme for heart failure**

## patients in clinical practice: the HerzMobil Tirol programme

Clin Res Cardiol 2022 Mar;111(3):294-307. doi: 10.1007/s00392-021-01912-0

### Overview:

The HerzMobil Tirol program is a multidimensional post-discharge disease management program for heart failure patients that incorporates a telemedical monitoring system and a comprehensive network of specialized heart failure nurses, resident physicians, and referral centers. In a non-randomized study of 508 acute heart failure patients, the program was found to be feasible and effective in reducing the primary endpoint of time to HF readmission and all-cause mortality within 6 months by 46% compared to usual care. The program also showed a reduction in the composite of recurrent HF hospitalization and death within 6 months and a lower mortality rate after 1 year.

### Detailed Description:

#### Goal

To evaluate the feasibility and effectiveness of a multidimensional post-discharge disease management program for heart failure patients using telemedical technology.

#### Methods

The study included 508 acute heart failure (AHF) patients managed by HerzMobil Tirol or usual care after discharge from hospital.

Patients in the HMT group underwent standardized evaluation, received structured follow-up, and were provided with a blood pressure and heart rate monitor, weighing scale, and smartphone for daily data acquisition.

Logistic regression models were used to calculate hazard ratios for primary outcomes including 6-month HF hospitalization and all-cause mortality at 6 months.

#### Findings

Management by HerzMobil Tirol was associated with a significant reduction (46%) in time to HF readmission within six months compared with usual care.

The composite endpoint was significantly lower with HMT than UC indicating that the program is effective.

Patients compliance among participants remaining in the HMT programme was high; only six out of two hundred fifty-one patients

were found negligent on data transfer but remained until completion after three months.

### Implications

Results suggest that specific disease management programs should be implemented widely following an acute heart failure event as they can prevent readmissions while improving clinical outcomes such as self-care behavior among patients.

Zou Y, Pesaranghader A, Song Z, Verma A, Buckeridge DL, Li Y

### Modeling electronic health record data using an end-to-end knowledge-graph-informed topic model

Sci Rep 2022 Oct 25;12(1):17868. doi: 10.1038/s41598-022-22956-w

### Overview:

The article describes a novel approach for modeling electronic health record (EHR) data using an end-to-end knowledge-graph-informed topic model called Graph ATtention-Embedded Topic Model (GAT-ETM). The GAT-ETM method leverages a medical knowledge graph to learn clinically meaningful disease topics from EHR data, and demonstrates superior performance over alternative methods for tasks such as drug imputation and disease diagnosis prediction. The approach has potential applications in computational phenotyping, patient stratification, and drug recommendations, and provides a promising avenue for refining disease phenotypes and discovering novel disease comorbidities from large-scale EHR datasets.

### Detailed Description:

#### Goal

The goal of the text is to present Graph ATtention-Embedded Topic Model (GAT-ETM), an end-to-end taxonomy-knowledge-graph-based multimodal embedded topic model that distills latent disease topics from EHR data by learning the embedding from a constructed medical knowledge graph.

#### Methods

GAT-ETM assumes a generative process for each patient in the EHR corpus, where topic mixture membership is drawn from logistic-normal and categorical distributions.

To extract meaningful and interpretable disease topics, a linear decoder is used to

reconstruct EHR data such that the linear projections can directly map to individual latent topics.

A graph augmentation strategy was proposed by connecting nodes with their ancestry nodes along taxonomy in order to maximize information flow among EHR nodes on the graph.

An ICD-ATC knowledge graph is leveraged to learn code embedding.

Empirically, it was found that mini-batch stochastic gradient descent worked well for updating models with large datasets.

### Findings

GAT-ETM demonstrated superior performance over alternative methods on all tasks including drug imputation and disease diagnosis prediction.

GAT-ETM learned clinically meaningful embeddings of codes and discovered accurate patient representations for stratification purposes as well as drug recommendations.

The study evaluated model performance using expert-derived rule-based labels across 12 chronic diseases; results showed high accuracy classification scores across all diseases tested.

### Implications

Incorporating multiple views through integrating knowledge graphs brings complementary information which improves phenotyping accuracy quantitatively & qualitatively.

Attention mechanism enables tracking input feature contributions which will be useful in understanding connections between different diseases.

<sup>1</sup> We are convinced that artificial intelligence in the clinical setting will become more and more important and helpful. So, this year we asked an artificial intelligence to create the content summaries for our best papers. We used the tool Explainpaper (<https://www.explainpaper.com/dashboard>, Subscription Plus Mode) and did the following:

- Marke Headline and EXPLAIN IN EXPERT STATUS using EXPLAIN MODEL MODEL GPT-4 (costs 19 \$ in addition)
- Use Explain COMPLETE PAPER using DETAILED SELECTION OF PAPER.

We present the unmodified original texts as supplied by the AI tool (only formatting was added afterwards).