OPTIMIZING CLINICAL PATHWAYS WITH FEDERATED DATA

Tim Müller/Rosanne J. Turner/Saba Amiri/Corinne Allaart/
L. Thomas van Binsbergen/Lea Dijksman/Tom van Engers/
Adam Belloum/Paola Grosso/Peter Grünwald/Karin Hoogoort/
Aki Härmä/Johanna Maria Hegeman/Jamila Alsayed Kassem/
Milen Kebede/Cees de Laat/Paul van der Nat/Auke Pals/
Floortje Scheepers/Sander Klous

Tim Müller, Scientific Programmer *
t.muller@uva.nl

Rosanne J. Turner, Clinical Data Science Researcher †
R.J.Turner@umcutrecht.nl

Saba Amiri, PhD Candidate *
s.amiri@uva.nl

Corinne Allaart, PhD Candidate, Vrije Universiteit, Computer Science Department
De Boelelaan 1105, 1081HV, Amsterdam, NL
c.allaart@antoniusziekenhuis.nl

L. Thomas van Binsbergen, Assistant Professor *
l.t.vanbinsbergen@uva.nl

Lea Dijksman, Senior Advisor/Researcher §
l.dijksman@antoniusziekenhuis.nl

Tom van Engers, Full Professor *
vanEngers@uva.nl

Adam Belloum, Assistant Professor *
a.s.z.belloum@uva.nl

Paola Grosso, Full Professor, *
p.grosso@uva.nl

Peter Grünwald, Professor, Centrum voor Wiskunde en Informatica, Machine Learning Group
Science Park 123, 1098XG Amsterdam, NL
Peter.Grunwald@cwi.nl

Karin Hoogoort, Lab Coordinator †
K.Hoogoort@umcutrecht.nl

Aki Härmä, Assistant Professor, Maastricht University, Department of Advanced Computing Sciences
Minderbroedersweg 4–6, 6211LK Maastricht, NL
aki.harma@maastrichtuniversity.nl

Johanna Maria Hegeman, St. Antonius Hospital, Psychiatry
Koekoekslaan 1, 3435CM Nieuwegein, NL
a.hegeman@antoniusziekenhuis.nl

Jamila Alsayed Kassem, PhD Candidate *
j.alsayedkassem@uva.nl

Milen Kebede, PhD Candidate *
m.g.kebede@uva.nl

Cees de Laat, Professor Emeritus *
C.T.A.M.deLaat@uva.nl

Paul van der Nat, Program Manager §
p.van.der.nat@antoniusziekenhuis.nl

Auke Pals, Senior Consultant, KPMG, Responsible AI
Laan van Langerhuize 1, 1186DS, Amstelveen, NL
pals.auke@kpmg.nl

Floortje Scheepers, Full Professor †
f.e.scheepers-2@umcutrecht.nl

Sander Klous, Professor *
s.klous@uva.nl

* University of Amsterdam, Informatics Institute (IvI)
Science Park 900, 1098XH Amsterdam, NL

† UMC Utrecht, Psychiatry
Heidelberglaan 100, 3584CX Utrecht, NL

§ St. Antonius Hospital, Department of Value-Based Healthcare
Koekoekslaan 1, 3435CM Nieuwegein, NL

Abstract: *Clinical pathways are currently difficult to optimize due to sensitive data that is typically distributed across organizations while rules and regulations constrain access and data processing. In this paper we describe a federated approach that can significantly reduce the efforts required to overcome these obstacles. First, we describe a standard conceptual workflow for optimizing clinical pathways, including all steps and involved stakeholders. This is followed by a translation of the workflow into a real-world scenario with an associated proof of principle to demonstrate how the scenario can be implemented on top of a federated framework. We present the most important results and conclude with an overview of the benefits for each of the stakeholders. Our most important outcomes are: the federated approach offers significant benefits for all relevant stakeholders and has little downsides. A policy-driven framework with embedded policy enforcement is crucial for successful adoption of a federated approach. Integration of safe statistics and synthetic data generation in a federated framework is straightforward and offers additional benefits, especially when setting up healthcare consortia. This solution is almost ready to be adopted by healthcare organizations as part of their regular operations.*

1. Introduction

Clinical pathways,[1] also known as integrated care pathways, aid organizations of care by promoting organized, collaborative care and shared decision-making for well-defined groups of patients based on evidence-based medicine.[2] One of the challenges in optimizing clinical pathways (for both efficiency of care and patient outcomes) in a patient-tailored manner is to get access to enough data that fit the characteristics of an individual patient because a) data are typically distributed across several organizations, potentially in different jurisdictions, and b) policies (privacy or other) typically restrict data access and processing, hampering clinical pathway optimization.

Many initiatives already exist to address these issues for specific situations. An example during the COVID crisis was the sharing of information between intensive care units for insight in capacity and quality purposes.[3] The standard workflow of such care path optimizations involves the collection of data and implementation of code within the computational environment of each site by data scientists, in an environment provided

---

1 KLOUS, S./WIELAARD, N. (2018). Building trust in a smart society: Managing in a modular, agile and Decentralized Way. Infinite Ideas.
2 ASKARI, M./TAM, J.L.Y.Y./AARNOUTSE, M.F./MEULENDIJK M. (2020). Perceived effectiveness of clinical pathway software: A before-re-after study in the Netherlands. Int J Med Inform., 2020.
3 GEUBBELS, E. L. P. E. (2023). The daily updated Dutch national database on COVID-19 epidemiology, vaccination and sewage surveillance. Sci Data, 469.

by IT officers. Data collection can be complex, due to legal requirements to either ask for patient consent and provide an option for data subjects to withdraw at any time, or to completely anonymize a dataset through case-specific and non-trivial means.[4] Furthermore, data transfer agreements and ethics approval must be arranged by management (e.g. department heads or principal investigator), privacy officers and ethics committees at each site to be compliant with rules and regulations. The required development of joined policies, operational guidelines and an interoperable (IT-)infrastructure under a concordant governance structure is a high effort and time-consuming process, which can take months or even years.[5, 6]

We propose a federated structure that can facilitate the described process and allows collaborating organizations to deal with policies and ethical considerations at a local level, limit dependencies between them and improve adaptability to new circumstances that would require new policies, enabling a consortium to control data access and data usage according to policies set by individual collaborating partners. These benefits can be best achieved when the solution is based on a federated approach with embedded policy enforcement.[7] The required infrastructures and workflows are configured and constrained based on policies specified by relevant stakeholders at various levels, e.g., privacy policies, security policies, access policies, usage policies, etc.[8] This novel approach largely eliminates the need to deploy specific IT infrastructure for each individual case as its configuration is managed by setting the appropriate policies for that case. Note that policies can be adjusted based on changes in external conditions as well (referred to as dynamic policy management). This can be especially relevant e.g. during crisis situations, avoiding the need for ad-hoc solutions. The approach is embedded in the solution via a policy enforcement mechanism avoiding circumventing policies without proper justification. The mechanism offers compliance, guaranteed with rules, regulations and other agreements, reducing the overhead of manual compliance monitoring and enforcement side by the responsible organization. The solution further features the generation of synthetic data, that can specifically be created for the analysis at hand, enabling the generalization of e.g., privacy related policies. In the rest of the paper, we will refer to the described solution, i.e., the architecture and its reference implementation as "the framework".[9, 10]

In this article, we will describe the results of a proof of principle to show the feasibility of this approach. The proof of principle is based on a simplified scenario extracted from a real-life use case in psychiatric treatments. The structure of this article is as follows. In section 2 we describe a typical workflow in a clinical setting with the individual roles and their expected contributions to the successful execution of an analysis aimed at optimizing a clinical pathway. Mapping out this process at a conceptual level from a clinical perspective is a crucial step. It reveals which stakeholders are involved and specifies their contribution. Based on these insights we can pinpoint the overhead the framework is able to reduce in an operational setting. The effects are twofold: 1) It reduces the effort required to optimize a clinical pathway, i.e., it optimizes the workflow and 2) It provides a means to improve the optimization of the pathway itself, e.g., by offering access to data that was previously unavailable. In section 3 we describe a real-life use case to make the conceptual description of section 2 tangible for a specific case. In section 4 we present the experimental setup and in section 5 the results of our proof of principle. Finally in section 6 we describe the benefits of our solution, including the incentives for the individual stakeholders and a way forward for adoption.

[4]   EMILY M WEITZENBOECK, PIERRE LISON, MALGORZATA CYNDECKA, MALCOLM LANGFORD (2022). The GDPR and unstructured data: is anonymization possible?, International Data Privacy Law, Volume 12, Issue 3, August 2022, Pages 184–206, https://doi.org/10.1093/idpl/ipac008.

[5]   TURNER, R.J. (2023). Safe Anytime-Valid Inference: from Theory to Implementation in Psychiatry Research. PhD thesis.

[6]   ALLAART, C. G./KEYSER, B./BAL, H./VAN HALTEREN, A. (2022, July). Vertical Split Learning-an exploration of predictive performance in medical and other use cases. In 2022 International Joint Conference on Neural Networks (IJCNN) (pp. 1–8). IEEE.

[7]   PARK, J./SANDHU R. (2002). Towards usage control models: beyond traditional access control. Proceedings of the seventh ACM symposium on Access control models and technologies.

[8]   UDDIN, M./SHAREEFUL, I./AL-NEMRAT, A. (2019). A dynamic access control model using authorising workflow and task-role-based access control. Ieee Access 7: 166676–166689.

[9]   VAN BINSBERGEN, L. T. et al. (2024). AMdEX Reference Architecture-version 1.0.0.

[10]   CUSHING, R. et al. (2021). Process data infrastructure and data services. Computers and informatics.

2. Conceptual workflow to optimize clinical pathways

In Fig. 1 a conceptual workflow deduced from a significant number of actual clinical pathways is presented.[11] In this section the context is provided for this workflow with the relevant stakeholders marked in **bold** font and the policies in *italic* font. A **physician** wants to optimize (either for efficiency and/or patient outcomes) the clinical pathway for a specific group of patients. It all starts with a research question that must be answered with a data analysis. Help is needed of a **data scientist** who can *translate desired medical insights into data analysis requirements and policies*. A **data privacy officer** *sets privacy related policies and guidelines* that need to be followed. Before the analysis can start, approval is needed from an **ethical committee** who *assesses the acceptability of the plan*. They *may have some conditions, e.g., related to the use of the analysis results, that need to be complied with as well*. Once approval is obtained, the **data scientist** can get access to the data and develop, train, and deploy a model that can run the analysis. The data scientist needs access to the IT infrastructure that meets the requirements and policies as defined in one of the previous steps. An **IT specialist** can provide access and *provide an infrastructure able to comply with and enforce the policies*. The **data scientist** needs to train the model using large amounts of data that either *require anonymization or consent for processing* by the **patient**. When the model is ready, the analysis can be run for individual **patients**. The **patients** *need to be informed about the data they need to provide and how it will be used, for example when identifiable data are used by the model or the model is updated with their data (note: data containing information about diagnosis and age are also identifiable). Their consent is required* as specified in their personal data access policies. If necessary, the **data scientist** may run or evaluate the analysis with the **physician** and the **patient**. They may also explain the results and answer any question. Throughout this workflow, **department heads** or **principal investigators** coordinate and prioritize all phases and *are responsible for compliance with policies*. They also manage the budget and funding and provide context and support.
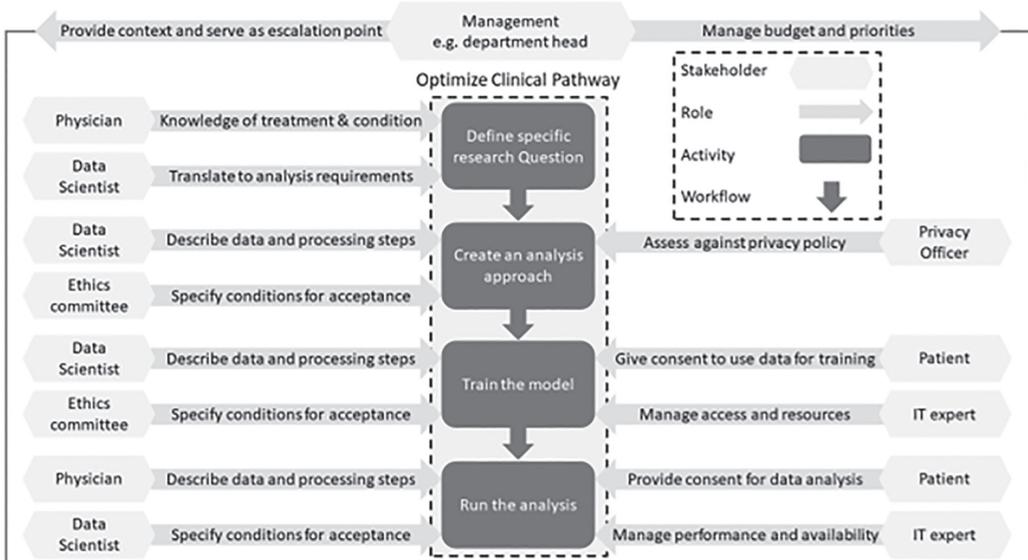
Fig. 1: Conceptual workflow of a single iteration in a continuous improvement process to optimize a clinical pathway.

---

[11] DANIELS, K./ROUPPE VAN DER VOORT, M. B./BIESMA, D. H./VAN DER NAT, P. B. (2022). Five years' experience with value-based quality improvement teams: the key factors to a successful implementation in hospital care. BMC Health Services Research, 22(1), 1271.

3. Description of real-life use case

Electroconvulsive therapy (ECT) is an intensive treatment that can be very effective in patients with severe depressive disorders. However, there are some severe potential cognitive side-effects and not all patients respond to ECT. Treatment outcome and the development of side effects are unfortunately not yet predictable on an individual level.[12] Potential predictors for ECT success are the presence of a personality disorder and having psychotic and/or catatonic traits. When someone has a personality disorder as a co-morbidity, the probability of responding to ECT decreases (odds ratio 0.3).[13] This is an important argument against starting ECT. On the other side, severe depression having certain traits (collection of symptoms), "catatonic" or "psychotic", increases the probability of response (odds ratio 1.7).[14] These traits may also have an interaction with personality disorder and change the probability of responding to ECT while having a personality disorder (giving rise to Simpson's paradox. A statistical phenomenon where the best performing organization may not have the highest success rate due to the fact that it may also treat the more complex cases). To obtain enough data for algorithms to deliver an accurate estimate of the probability that the patient responds to treatment, data needs to be combined across multiple hospitals. In this proof of principle, we combined data from UMC Utrecht and St. Antonius hospital to train algorithms that give guidance for this particular problem. Retrospective, routinely collected data on the traits described above and treatment outcome from electronic patient files from patients undergoing ECT at both sites from 2009–2019 were used (248 and 103 complete cases, respectively at each site, more details on data collection can be found in[15]. **Patients** must give their consent for their data being shared with the other hospitals for the study, and data transfer agreements and ethics approval must be arranged by **management (principal investigator)**, **privacy officers** and **ethics committees** at each site. This is a high effort and time-consuming process. In our experience it takes months or even years.[16, 17]

4. Experimental setup

We now present two scenarios to optimize the workflow for collaborating with data across hospitals to improve clinical care pathways, with were both tested on part of the ECT clinical care pathway at the two participating hospitals as a proof of principle. Both scenarios make use of a framework[18] able to perform an analysis on distributed datasets. The unique feature of this framework is its federation, i.e., participating organizations maintain their full autonomy, e.g., in providing access to data and enforcing policies.

---

[12] Van der Does, Y./Turner, R. J./Bartels, M. J./Hagoort, K./Metselaar, A./Scheepers, F./Grünwald P. D./Somers M./van Dellen, E. (2023). Outcome prediction of electroconvulsive therapy for depression. Psychiatry Research, 326, 115328.

[13] Prudic, J./Olfson, M./Marcus, S. C./Fuller, R. B./Sackeim, H. A. (2004). Effectiveness of electroconvulsive therapy in community settings. Biological psychiatry, 55(3), 301–312.

[14] Van Diermen, L./Van Den Ameele, S./Kamperman, A. M./Sabbe, B. C./Vermeulen, T./Schrijvers, D./Birkenhäger, T. K. (2018). Prediction of electroconvulsive therapy response and remission in major depression: meta-analysis. The British journal of psychiatry, 212(2), 71–80.

[15] Van der Does, Y./Turner, R. J./Bartels, M. J./Hagoort, K./Metselaar, A./Scheepers, F./Grünwald P. D./Somers M./van Dellen, E. (2023). Outcome prediction of electroconvulsive therapy for depression. Psychiatry Research, 326, 115328.

[16] Turner, R.J. (2023). Safe Anytime-Valid Inference: from Theory to Implementation in Psychiatry Research. PhD thesis.

[17] Allaart, C. G./Keyser, B./Bal, H./Van Halteren, A. (2022, July). Vertical Split Learning-an exploration of predictive performance in medical and other use cases. In 2022 International Joint Conference on Neural Networks (IJCNN) (pp. 1–8). IEEE.

[18] Alsayed Kassem, J./Müller, T./Esterhuyse, C. A./Kebede, M. (2023a). The EPI Framework: A data Privacy by Design framework to support healthcare use-cases. In Future Generation Computer Systems.

4.1.   Framework, workflows and policies

A proof of principle was set up on infrastructure owned by SURF, the Utrecht research medical center and the St. Antonius Hospital. The latter two domains each have a partition of the ECT dataset and offered infrastructure to perform local processing on their own data. SURF acted as a trusted third-party, hosting any federal components of the framework and offering only computational infrastructure. SURF is trusted by both hospitals to aggregate their local results into a global result and store it. The infrastructure is based upon our generic highly flexible data-sharing framework that can be configured to meet and enforce a wide variety of policies. The analysis was implemented in the framework by separating it into various tasks as specified in a computational workflow (see Fig. 2).[19]
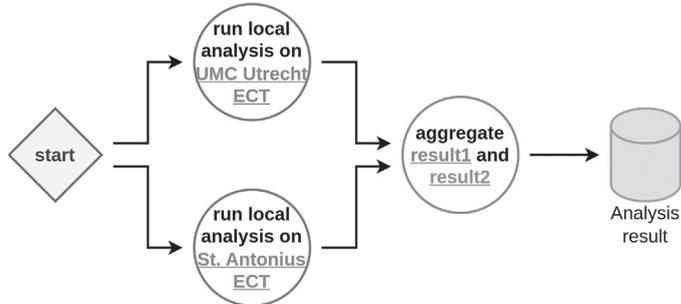
Fig. 2: Formal representation of workflows for the Proof-of-Principle (PoP). This figure shows that the work can be segmented into tasks that can be executed on different sites. The framework takes care to move data between sites as necessary, while checking in with the defined policy to keep the transfers compliant to the data-access restrictions.

During the execution of the workflow, the federal part of the framework orchestrates the execution of individual tasks by requesting domains to do so and to transfer data when necessary for the execution. At every request, the local delegates of the framework consult with local policy reasoners programmed with the respective domain's policy (encoding them in e.g. the eFLINT[20] policy language). The reasoner can determine whether the request is compliant, and only if it does will the local delegate fulfill the request. Otherwise, they refuse and workflow execution halts, preventing each domain and its data from participating in workflows which would violate their own policy.[21]

Aside from enforcing normative policies about task execution and data access, low-level, networking policies that enforce security and privacy requirements[22] are also supported. In particular, by leveraging Virtualized Network Functions (VNF), the system can create security solutions like firewalls, encryption, and/or Intrusion Detection Systems (IDS) on-demand. This significantly reduces the overhead for IT officers, as cumbersome negotiations about setting up Virtual Private Networks (VPNs) are replaced by specifying a network policy.

In the scenario, the reasoners are programmed with policies that limit any execution done by domains to only code part of the use-case; the hospitals do not allow transfers of any of their data except of the results of the local analysis; and the trusted third-party only allows researchers to download the final, aggregated result.[23]

---

[19]   Equivalent workflow scripts can be found here: https://github.com/epi-project/proof-of-principle.

[20]   VAN BINSBERGEN, L. T./LIU, L-C./VAN DOESBURG, R./VAN ENGERS, T. (2020). eFLINT: a domain-specific language for executable norm specifications. In Proceedings of the 19th ACM SIGPLAN International Conference on Generative Programming: Concepts and Experiences (GPCE 2020). Association for Computing Machinery, New York, NY, USA (pp. 124–136). doi: 10.1145/3425898.3426958.

[21]   ESTERHUYSE, C. A./MÜLLER, T./VAN BINSBERGEN, L. T./BELLOUM, A. S. Z. (2022), Exploring the Enforcement of Private, Dynamic Policies on Medical Workflow Execution. In 2022 IEEE 18th International Conference on e-Science (e-Science), Salt Lake City, UT, USA (pp. 481–486), doi: 10.1109/eScience55777.2022.00086.

[22]   ALSAYED KASSEM, J./ZHONG, L./TAAL, A./GROSSO, P. (2023b). Adaptive services function chain orchestration for digital health twin use cases: Heuristic-boosted Q-learning approach. On arXiv. doi: 10.48550/arXiv.2304.12853.

[23]   Equivalent eFLINT policy can be found here: https://github.com/epi-project/proof-of-principle.

4.2. Federated Direct Access and Learning

In the scenario described in section A, the data scientist assisting the psychiatrist negotiates with the participating hospitals to use the required data for this specific analysis. Participating hospitals are willing to grant the data scientist permission to perform a specific task given that the data stays under their control, no direct access to this data was required by the researchers, and only the approved analysis can be executed. The data scientist only gets access to a statistical representation of the local model, which guarantees the privacy of the patients involved. The framework supports privacy preserving federated learning methods to further enhance privacy protection, although this feature is not used in the first scenario. The first scenario specifically offers a transparent, exact approach to estimate the effect of personality disorder on the response probability after ECT, correcting for catatonic and psychotic traits. Even though the first scenario makes previously unfeasible analyses possible, negotiations about access conditions and implementing the required controls can be cumbersome, and federated versions of machine learning models are not suitable for every use-case due to issues such as availability or non-iid data distribution.[24] This is why we also offer a second scenario that includes an additional step to generate synthetic data based on the source data required for the analysis (see the workflow in Fig. 3). Synthetic data guarantees the privacy of patients involved even if control over that data would be lost, adding an extra layer of protection. Policies for generating synthetic data can therefore be generalized, simplifying negotiations to a level that allows highly automated processing. This reduces the overhead on the compliance side of the organization, as described in the introduction and can have a game changing impact on how physicians and data scientist's work.
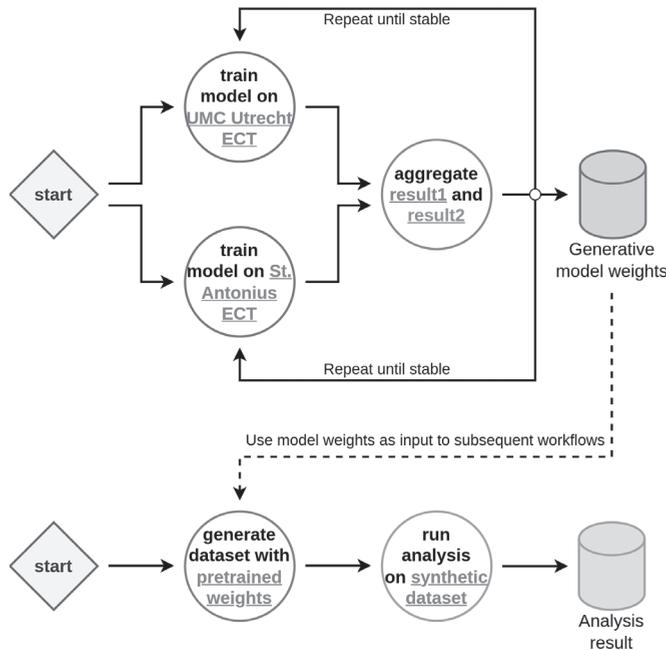

Fig. 3: Workflows implementing the second scenario with synthetic data. Here, we can first train a synthetic generation algorithm in a first workflow (federated), generating weights. Then, instead of performing the analysis on the ECT dataset in a federated manner, it can be performed directly on newly generated data using the pre-trained weights.

---

24   Amiri, S./Belloum, A./Nalisnick, E./Klous, S./Gommans, L. (2022, June). On the impact of non-IID data on the performance and fairness of differentially private federated learning. In 2022 52nd Annual IEEE/IFIP International Conference on Dependable Systems and Networks Workshops (DSN-W) (pp. 52–58). IEEE.

5.   Results

The direct federated analysis resulted in two anytime-valid confidence sequences based on the data from the two hospitals combined, one for each stratification level: without catatonic or psychotic traits ("no traits" in Fig. 4), or with at least one of those ("cat./psych. traits" in Fig. 4). Due to the anytime-validity of the algorithm researchers could for example have a dashboard running where they can follow the confidence intervals getting narrower as the study progresses. At any time during the study, they can look at the most recent confidence interval (one of the vertical lines in Fig. 4) and observe the exact 95% confidence interval for the effect of personality disorder on the probability of response to ECT. This enables them to decide on the best treatment much faster compared to when a classical meta-analysis set up would be used: as can be observed in Fig. 4, bottom panel, at the end of the study, it has become clear that having a personality disorder without catatonic or psychotic traits has a negative effect on the probability of responding to ECT, with a confidence interval of the effect ranging from –0.08 to –0.42 (the most-right vertical bar in the lower panel of Fig. 4). In Fig. 4, it can be observed that 0 is no longer included in the confidence interval after including 20 patients, so the negative effect of having a personality disorder without traits on ECT outcome was already clear early during the study. This allows physicians to take the decision to not recommend ECT to these patients a lot earlier. In a classical study setup, for example a combination of clinical trials at multiple hospitals, one must wait until data collection is finished at each individual hospital before results can be combined between hospitals (formal comparisons of classical (statistical and meta-) analysis and anytime-valid analysis can be found in[25] and[26]).

The comparative analysis of the classifiers' performance on real and synthetic data presents compelling insights (see Fig. 5). The results with synthetic data, demonstrate superior performance, acquiring an accuracy of 0.333 and an F1 score of 0.417. Contrastingly, classifiers trained on real data exhibit performances of an accuracy of 0.238 and F1 score of 0.385. The superiority of synthetic data compared to real data for the specific classification task can be attributed to the generation process mitigating some of the issues commonly associated with real-world medical data, e.g. class imbalance, missing values, and noise.
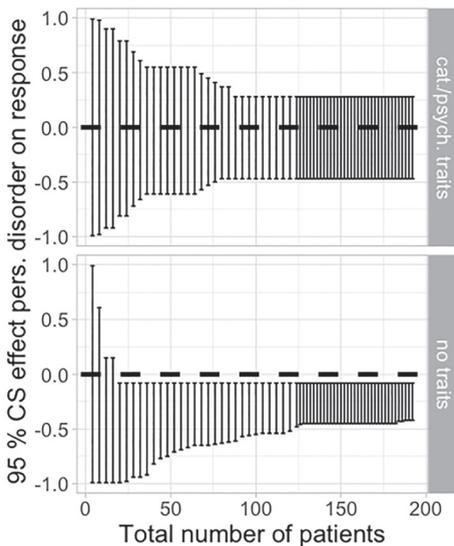
Fig. 4: The stratified 95% confidence sequences (collection of confidence intervals over time during a study) resulting from the direct federated analysis with anytime-valid confidence intervals of the ECT dataset. The confidence sequence concerns the effect of personality disorder on the probability of responding to ECT, stratified according to the existence of psychotic and/or catatonic traits.

[25]   Turner, R.J./Grunwald, P.D. (2023, April). Safe Sequential Testing and Effect Estimation in Stratified Count Data. In International Conference on Artificial Intelligence and Statistics (pp. 4880–4893). PMLR.

[26]   Ter Schure, J./Grünwald, P.D. (2019). Accumulation Bias in meta-analysis: the need to consider time in error control. F1000Research, 8.
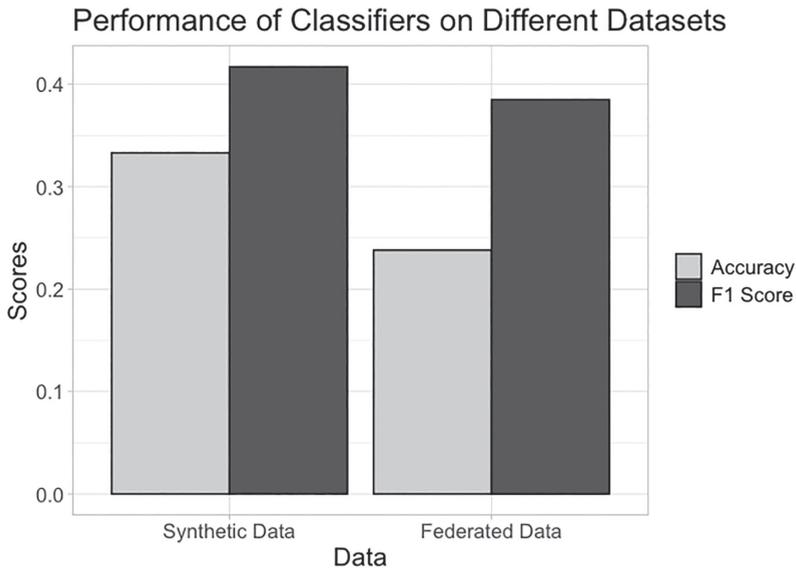
Fig. 5: Comparative Performance Metrics of Classifiers Trained on Real and Synthetic Datasets. This bar chart illustrates the performance of classifiers trained on synthetic data and real data (compromising joint datasets from St. Antonius and UMCU). Performance is measured via two metrics: accuracy and F1-score. The classifiers trained on synthetic data demonstrate superior performance.

The addition of synthetic data generation to the analysis chain does add a layer of complexity that has to be taken into account. Besides the potential overhead of qualitative evaluation, synthetic data can either omit or include features with respect to the original data. The advantages and disadvantages of synthetic data should therefore be considered carefully for each case.

6. Conclusion

In this article we demonstrated the technical and statistical components to optimize clinical pathways with integrated data from multiple organizations at scale in an operational environment. Although this proof of principle of a federated approach contains a simplified version of reality, it has been applied to a real-life psychiatric use case to optimize the decision about starting an ECT treatment. The results show that the decision about the best treatment can be made much faster compared to when a classical meta-analysis set up would be used. We added the feature of synthetic data generation to the system. This reduces the challenges related to cross-domain (data) governance. Our framework allows to generate synthetic data "on-the-fly", i.e., only suitable for the analysis at hand. This means only the correlations between the variables required for a specific analysis are taken into account when the synthetic data is generated, avoiding the overly complex task of generating a fully representative synthetic data set suitable for all purposes. The results show that the performance of the analysis on synthetic data in our proof of principle experiment is superior compared to that of an analysis on real data. Synthetic data enables healthcare research to meet scientific standards as most journals require the publication of source data to allow for validation of the results, that can be problematic due to privacy related restrictions.

The federated approach as implemented by our framework increases flexibility and reduces the overhead when collaborations are set up. Each participating party will individually benefit from the introduction of this framework and when more organizations adopt the framework, additional benefits will appear due to easier alignment on and adoption of protocols and policies (i.e., standardization) and the resulting further reduction of overheads. Below we summarize the benefits of the system for the most common stakeholders in a wide variety of healthcare use cases:

Data Scientist: some additional effort is needed to adopt methods suitable to make optimal use of the framework. Anytime-valid confidence intervals are an example of such a method, synthetic data is another example. Both methods already provide benefits when they are applied within a single organization, e.g., the ability to implement a more iterative approach. On top of that, they significantly reduce the complexity when the number of participating organizations increases.

Physician: optimizing clinical pathways across multiple organizations for individual patients is currently a cumbersome if not prohibitively time-consuming task. After a research question has been defined, it typically requires dedicated arrangements with other hospitals to deal with organizational and infrastructural constraints. The framework introduced in this article simplifies and standardizes the implementation of these constraints, allowing the physician (and the data scientist) to spend more time on answering the research question.

Medical Ethical Committee and Data Privacy Officer: the diversity in research questions and in data, especially when combined, makes risk management a complex and time-consuming task. On top of that, the limitations arising from the resulting constraints and policies do not fit well with the preferred agile way of working when a model is developed. The framework can greatly simplify several activities performed by the ethical committee with its policy standardization and automation mechanisms, especially in combination with synthetic data. This is beneficial for single organization use cases, but the impact will be even larger when an analysis extends over multiple organizations.

IT specialist: the main benefit of the framework from an IT operations perspective is the interaction between the policy management layer and the infrastructure layer. The system can automatically be reconfigured based on e.g., network or security policies by applying software defined infrastructure techniques. As a result, a generic infrastructure layer is suitable to address a wide variety of use cases. The infrastructure is configured by specifying the desired policies, which is much more efficient than setting up an environment from scratch for each request, especially in cross-domain situations.

Department head and/or Principal Investigator: the stakeholder that arguably benefits the most from reduced overheads and a more agile way of working is the department head or principal investigator. Due to the increased flexibility and reduced overhead, the investments (in time and material) to establish new collaborations will be lower. Hence, a way of working can arise where consortia are established and terminated in a more agile fashion. In other words, the impact of a "wrong" decision is reduced, making it possible to adopt a trial & error or learning by doing approach when setting up new consortia. Making agreements and allocating resources becomes more efficient, because if it doesn't work, it is easily changed and if it fails completely, losses are limited.

Patient: the real winner of all the above is the patient. Better privacy protection, faster responses, and optimization of an entire clinical pathway instead of just a segment are all within reach with the functionality offered by this framework.

More generic to the particular benefits for the clinical pathways use case are the federated data sharing infrastructure governance and normative pluralism. Many domains require IT-solutions that involve multiple stakeholders, multiple data sets and data processing requirements, and multiple regulatory demands (directives, laws and business policies) that have to be complied to for legitimate data processing. The infrastructure set up for the clinical pathway use case is set up based on general principles and generic architecture that is equally useful to other large-scale industrial applications. Examples of such use cases are amongst others banking

organizations that, while being competitors, have to collaborate in order to fulfill their obligations set down in anti-money laundering regulations. Another example is collaborations between stakeholders in logistics, that have to follow import- and export-regulations, have reporting obligations to governmental institutions, but want to keep their transaction data confidential in order to protect themselves against competitors. In fact, the clinical pathway use case serves as a framework for any multi-party, cross-jurisdiction use case that is based on a federated governance model, respecting the individuality and sovereignty of stakeholders and protecting the interest of the data owners and data subjects alike.

7. Future work

Although our Proof of Principle was a success, a lot of work still must be done before this framework is fully operational. Currently we are working on a Target Operating Model to work out the implementation details of this solution. An important part of that model is the audit function. As the framework is policy driven and policy enforcement is embedded, it can be considered as "Compliant by Design". Nevertheless, independent and highly automated third party verification will be required to adhere to proper risk management procedures, following good practices of the three lines of defense model.[27] On top of that, the business case must become more concrete by making a comprehensive cost versus benefits analysis, including an estimate of how much time it will take to setup an analysis or to train a machine learning model. The design and validation of the ML models and experiments can be extended to more data, more complicated models and experimental designs which will be enabled by the vast literature and implementations already available in the machine learning field. The software developed for this proof of principle must become production ready. Fortunately, this software is not healthcare specific and several projects in other domains have already been funded to complete this task. In the meantime, we are working on a project proposal that will allow healthcare organizations to adopt federated clinical pathway optimizations with this framework into their operational activities. The proposal will include training and implementation aspects required for adoption of the framework by healthcare professionals. This is needed to make sure adoption of this framework in the healthcare sector is prioritized.

8. References

Allaart, C. G./Keyser, B./Bal, H./Van Halteren, A. (2022, July). Vertical Split Learning-an exploration of predictive performance in medical and other use cases. In 2022 International Joint Conference on Neural Networks (IJCNN) (pp. 1–8). IEEE.

Alsayed Kassem, J./Müller, T./Esterhuyse, C. A./Kebede, M. (2023a). The EPI Framework: A data Privacy by Design framework to support healthcare use-cases. In Future Generation Computer Systems.

Alsayed Kassem, J./Zhong, L./Taal, A./Grosso, P. (2023b). Adaptive services function chain orchestration for digital health twin use cases: Heuristic-boosted Q-learning approach. On arXiv. doi: 10.48550/arXiv.2304.12853.

Amiri, S./Belloum, A./Nalisnick, E./Klous, S./Gommans, L. (2022, June). On the impact of non-IID data on the performance and fairness of differentially private federated learning. In 2022 52nd Annual IEEE/IFIP International Conference on Dependable Systems and Networks Workshops (DSN-W) (pp. 52–58). IEEE.

Askari, M./Tam, J.L.Y.Y./Aarnoutse, M.F./Meulendijk M. (2020). Perceived effectiveness of clinical pathway software: A before-after study in the Netherlands. Int J Med Inform., 2020.

Aspland, E./Gartner, D./Harper, P. (2019). Clinical pathway modelling: a literature review. Health Systems, 10(1), 1–23.

Van Binsbergen, L. T. et al. (2024). AMdEX Reference Architecture-version 1.0.0.

---

[27] Klous, S./Wielaard, N. (2018). Building trust in a smart society: Managing in a modular, agile and Decentralized Way. Infinite Ideas.

van Binsbergen, L. T./Liu, L-C./van Doesburg, R./van Engers, T. (2020). eFLINT: a domain-specific language for executable norm specifications. In Proceedings of the 19th ACM SIGPLAN International Conference on Generative Programming: Concepts and Experiences (GPCE 2020). Association for Computing Machinery, New York, NY, USA (pp. 124–136). doi: 10.1145/3425898.3426958.

Cushing, R. et al. (2021). Process data infrastructure and data services. Computers and informatics.

Geubbels, E. L. P. E. (2023). The daily updated Dutch national database on COVID-19 epidemiology, vaccination and sewage surveillance. Sci Data, 469.

Daniels, K./Rouppe van der Voort, M. B./Biesma, D. H./van der Nat, P. B. (2022). Five years' experience with value-based quality improvement teams: the key factors to a successful implementation in hospital care. BMC Health Services Research, 22(1), 1271.

Van Diermen, L./Van Den Ameele, S./Kamperman, A. M./Sabbe, B. C./Vermeulen, T./Schrijvers, D./Birkenhäger, T. K. (2018). Prediction of electroconvulsive therapy response and remission in major depression: meta-analysis. The British journal of psychiatry, 212(2), 71–80.

van der Does, Y./Turner, R. J./Bartels, M. J./Hagoort, K./Metselaar, A./Scheepers, F./Grünwald P. D./Somers M./van Dellen, E. (2023). Outcome prediction of electroconvulsive therapy for depression. Psychiatry Research, 326, 115328.

Dwork, C./Roth, A. (2014). The algorithmic foundations of differential privacy. Foundations and Trends® in Theoretical Computer Science, 9(3–4), 211–407.

Dyda, A. et al. "Differential privacy for public health data: An innovative tool to optimize information sharing while protecting data confidentiality." Patterns 2.12 (2021).

Emily M Weitzenboeck, Pierre Lison, Malgorzata Cyndecka, Malcolm Langford (2022). The GDPR and unstructured data: is anonymization possible?, International Data Privacy Law, Volume 12, Issue 3, August 2022, Pages 184–206, https://doi.org/10.1093/idpl/ipac008.

Esterhuyse, C. A./Müller, T./van Binsbergen, L. T./Belloum, A. S. Z. (2022), Exploring the Enforcement of Private, Dynamic Policies on Medical Workflow Execution. In 2022 IEEE 18th International Conference on e-Science (e-Science), Salt Lake City, UT, USA (pp. 481–486), doi: 10.1109/eScience55777.2022.00086.

Menger, V./Spruit, M./Hagoort, K./Scheepers, F. (2016). Transitioning to a data driven mental health practice: collaborative expert sessions for knowledge and hypothesis finding. Computational and mathematical methods in medicine.

Klous, S./Wielaard, N. (2018). Building trust in a smart society: Managing in a modular, agile and Decentralized Way. Infinite Ideas.

El Ouadrhiri, A/Abdelhadi, A. "Differential privacy for deep and federated learning: A survey." IEEE access 10 (2022): 22359–22380.

Park, J./Sandhu R. (2002). Towards usage control models: beyond traditional access control. Proceedings of the seventh ACM symposium on Access control models and technologies.

Prudic, J./Olfson, M./Marcus, S. C./Fuller, R. B./Sackeim, H. A. (2004). Effectiveness of electroconvulsive therapy in community settings. Biological psychiatry, 55(3), 301–312.

Ter Schure, J./Grünwald, P.D. (2019). Accumulation Bias in meta-analysis: the need to consider time in error control. F1000Research, 8.

Turner, R.J. (2023). Safe Anytime-Valid Inference: from Theory to Implementation in Psychiatry Research. PhD thesis.

Turner, R.J./Grunwald, P.D. (2023, April). Safe Sequential Testing and Effect Estimation in Stratified Count Data. In International Conference on Artificial Intelligence and Statistics (pp. 4880–4893). PMLR.

Uddin, M./Shareeful, I./Al-Nemrat, A. (2019). A dynamic access control model using authorising workflow and task-role-based access control. Ieee Access 7: 166676–166689.