# Discovery of logic-probabilistic rules from COVID-19 vaccine antibody response in older people: results from the GeroCovid VAX Study

Michela Vespa[1](✉)[0009−0004−4350−8151], Francesca Remelli[2][1111−2222−3333−4444], Elena Bellodi[1][0000−0002−3717−3779], and Raffaele Antonelli Incalzi[3][0000−0003−2100−2075]*

[1] Università di Ferrara, Dipartimento di Ingegneria
[2] Università di Ferrara, Dipartimento di Scienze mediche
[3] Università Campus Bio-Medico di Roma

**Abstract.** This paper presents a novel approach to handle uncertain, noisy and incomplete biomedical data through Probabilistic Logic Programming. Unlike previous systems, based on Inductive Logic Programming (ILP) for rule discovery or on user-defined probabilities, our approach enables the automatic discovery of probabilistic integrity constraints directly from biomedical data. These constraints are annotated with a probability and can assign a probability of belonging to the positive class to logical interpretations, which are used to formalize patients' medical records. We apply the PASCAL ("ProbAbiliStic inductive ConstrAint Logic") algorithm for learning such constraints, and we show that it outperforms traditional ILP and data mining approaches, such as TILDE, Aleph and Association Rules, on the data relative to long-term care facility residents participating in the GeroCovid Vax study.

**Keywords:** Probabilistic Logic Programming · integrity constraints · Biomedical Data Mining · COVID-19 vaccine.

## 1 Introduction

Medical data often contains uncertainty due to factors like biological differences, diverse disease progressions, measurement errors, and incomplete records. This uncertainty complicates efforts to derive reliable insights, as traditional analytical methods often produce oversimplified models that fail to capture the probabilistic relationships in the data. Logic-based methods provide an interpretable approach to this task, enabling the representation of relationships as human-readable rules in the form of First Order Logic (FOL) clauses (e.g., $\forall(X)\ Hypertension(X) \rightarrow PoorVaccineResponse(X)$). In particular, Inductive Logic Programming (ILP) systems automate the discovery of such rules directly from data. However, these systems generate deterministic rules, which fail to capture the inherent uncertainties present in real-world datasets. To address this

*on behalf of the GeroCovid Vax Working Group

limitation, Probabilistic Logic Programming (PLP) extends the capabilities of ILP by integrating probability theory with logical representations, enabling the expression of uncertain relationships while retaining the interpretability of symbolic Artificial Intelligence (AI). In this work, we apply PASCAL ("ProbAbiliStic inductive ConstrAint Logic") [15], an algorithm that learns logic-probabilistic theories, to the data collected in the GeroCovid VAX Study [20].

The COVID-19 pandemic has generated a lot of observational healthcare data, and this has highlighted the importance of data-driven approaches to help in understanding the complex interplay between patients' characteristics and clinical trajectories after the acute disease. The GeroCovid VAX Study emerged as a significant multicenter longitudinal study during the pandemic, focusing on understanding the clinical characteristics and outcomes of COVID-19 in older people, resident in long-term care facilities, who had undergone at least one dose of any type of anti-SARS-CoV-2 vaccination. In a subsample group, immune response at timed intervals from the first dose was investigated, aiming to identify factors influencing vaccine efficacy.

We compare PASCAL against established ILP methods such as TILDE [3] and Aleph [18], and traditional Association Rule Mining (ARM) algorithms like FP-Growth [10] and Eclat [23]. In order to apply PASCAL, we formalize the medical information of each patient as a labeled (positive/negative) interpretation. The theories learned by PASCAL are composed of probabilistic integrity constraints which assign a probability of being positive to new interpretations. TILDE and Aleph were chosen as they are among the most representative for the learning from interpretation setting in ILP (where the goal is to learn a theory that covers the positive interpretations and rules out the negative ones), while FP-Growth and Eclat can efficiently mine ARs. We demonstrate that PASCAL outperforms them according to several machine learning performance metrics and the clinical relevance of the findings. Models learned by PASCAL prove particularly effective in addressing the GeroCovid dataset's challenges, including missing data, heterogeneous clinical profiles, and a limited sample size.

The paper is organized as follows: Section 2 reviews related work. Section 3 introduces our methodology for representing clinical data using FOL. Section 4 describes the algorithms employed in the experiments. Section 5 presents experimental results on the GeroCovid data. Section 6 concludes this work.

## 2 Related Work

Medical data presents unique challenges for automated analysis due to its inherent complexity and heterogeneity. These challenges are compounded by uncertainty from missing data, variable patient responses, and the retrospective nature of healthcare records, which suffer from inconsistent collection patterns and varying documentation practices across providers and time periods [12]. The application of Logic Programming (LP) and ARM methodologies in biomedical domains has been an active area of research, yet challenges persist in developing systems that can effectively handle uncertainty while maintaining interpretabil-

ity. LP enables the representation of medical knowledge through *logical rules*. An example of this approach can be found in [11], which develop a decision system using Prolog-based argumentation logic to formalize clinical reasoning patterns into executable rules. In [14] ILP techniques were used to discover disease associations and risk factors for cholangitis and breast angiosarcomas from pathology data. Rule mining techniques, such as ARs, decision-tree-based methods and Learning Classifier Systems [21], are widely used in biomedical data analysis to uncover clinical patterns [2, 8, 19] with Apriori and Apriori-like algorithms. [7] apply fuzzy sets to address imprecision in medical data. While fuzzy theory improves the interpretability of qualitative data, it lacks a quantitative framework based on probability theory for uncertainty.

PLP has emerged as a promising field for medical data analysis due to its ability to combine logic and probability to handle uncertainty. Prior research [4] has applied PLP to perform breast cancer diagnosis and predict malignancy through the SkILL system [5]. SkILL uses TopLog to learn candidate rules and the ProbLog Yap library as a probabilistic inference engine. Both examples for the target predicate and the background knowledge (BK) are manually annotated with probabilities based on medical knowledge. Instead, our work focuses on *learning* logic-probabilistic constraint theories directly from data. [22] employ the principle of maximum entropy to build a knowledge base for analyzing brain tumor data, combining expert knowledge and statistical data to handle uncertain information, which can then be queried for diagnosis and prognosis. [16] apply Markov Logic Networks for both learning and inference on genetic data, combining FOL with probabilistic graphical models to capture complex interactions. While this method effectively integrates domain knowledge and handles uncertainty, it can become difficult to interpret as the networks grow; in contrast, our models are highly interpretable.

## 3 Clinical Data Representation

### 3.1 First-Order Logic

In order to apply PASCAL, TILDE and Aleph we formalized clinical data into logical interpretations. Each interpretation corresponds to a patient and consists of a set of ground facts whose predicates represent a patient's attributes and their arguments the attribute values. Missing data was handled by excluding the relative attribute from the interpretation. The attributes include the number of COVID-19 vaccine doses received (from 1 to 3), mobility levels, sex, age, physical activity over the last 3 months, the total number of medications taken, with polypharmacy in case of more than 5 medications. Chronic diseases, such as frailty, chronic kidney disease, poor nutrition, cancer, immune disease, inflammatory bowel disease, liver disorders, diabetes, obesity, osteoarthrosis, arterial hypertension, cardiovascular diseases, atrial fibrillation, peripheral arterial diseases, chronic respiratory diseases, depression, anxiety, cognitive diseases, epilepsy, Parkinson, thyroid diseases, heart ischemic diseases, urological diseases

were also included, along with the total count of chronic conditions. The humoral immune response was measured before vaccination (T0), 2 months after the first dose (T1), 6 months after the first dose (T2), and 12 months after the first dose (T3). Changes in antibody levels were also calculated to assess antibody response over time (from T0 to T1, T1 to T2, and T2 to T3). Information on COVID-19 infections prior to vaccination, at six and twelve months was included. According to the World Health Organization (WHO) clinical scale, different patient statuses were registered at T0, T1, T2, T3, ranging from no infection to severe illness requiring hospitalization. Patient outcomes were classified as: clinical improvement, no significant changes, significant adverse events, death, transfer to another facility or loss to follow-up. Additionally, vaccination status for COVID-19, flu, pneumococcal disease, and herpes zoster was recorded.

In our mapping the first argument is the patient's ID. The following interpretation represents a subject identified as 'it_0044_003650':

```
mobility_level(it_0044_003650,maximum_assistance).
frailty(it_0044_003650).
antibody_response_0(it_0044_003650,4.81).
outcome(it_0044_003650,no_major_change).
covidvaccine(it_0044_003650,bnt162b2,2_doses).
fluvaccine(it_0044_003650,No).
patient(it_0044_003650,pos).
```

The patient requires maximum assistance in moving and is frail. At T0 anti-COVID antibodies were equal to 4.81 BAU/ml. The patient remained stable throughout the observation period with no major changes. She received 2 doses of the Pfizer-BioNTech (BNT162b2) COVID-19 vaccine but not the flu one. As TILDE, Aleph and PASCAL are supervised learning algorithms, we had to label interpretations as positive or negative, according to one of the following two criteria: 1) patients were classified as positive if at least one antibody response at different times T was $\geq 33.8$ or if it was $\geq 33.8$ at all time points, and classified as negative in the other cases; 2) mortality: patients were labeled as positive if they died during the study. Note that the last ground fact is always defined by a `patient/2` predicate to specify the class (pos/neg) of interpretations.

### 3.2 Association Rules

Let $\mathcal{I}$ be a set of items, where each item is a literal *attribute=value*. A transaction $T \subseteq \mathcal{I}$ is a subset of $\mathcal{I}$. Association rules are defined on a set of transactions $\mathcal{T}$. An AR $R$ is in the form $X \rightarrow Y$, where $X$ (antecedent) and $Y$ (consequent) are two sets of items. Support and confidence are the two main parameters used to find relationships between items [1]. The transformation of clinical data for ARM involved several preprocessing steps. Each patient corresponds to a transaction and each of her attributes to an item. Missing values in the dataset were excluded from the corresponding transactions. Attributes with categorical values (e.g., sex, vaccination type) were converted using One-Hot Encoding: each categorical value is converted into one binary attribute. Continuous attributes (e.g., antibody

levels) were discretized into meaningful ranges or into 2 groups with respect to a threshold (low_antibody_response, high_antibody_response), based on medical expert input. These ranges were then encoded into binary attributes. Finally, each transaction lists only the binary attributes taking on value 1. The patient previously described by a logical interpretation becomes:

```
TID  Itemset
1    {it_0044_003650, maximum_assistance, frailty, low_antibody_response,
      no_major_change, bnt162b2, 2_doses, no_flu_vaccine}
```

where TID is the Transaction ID and Itemset is the set of items describing the patient. As FP-Growth and Eclat are unsupervised learning algorithms, we did not need to label the patients.

## 4 Rule Learning

In the following, we describe how the algorithms used for learning (P)LP theories and AR work.

### 4.1 Learning (probabilistic) logic models

TILDE[4] is an ILP algorithm that constructs logical decision trees (DT) from data under the learning from interpretations setting. It extends traditional DT learning algorithms, which typically operate on propositional data, to handle relational data. TILDE learns a DT by recursively partitioning the data using FOL queries, guided by a heuristic. It generates queries via refinement operators, leveraging a knowledge base (KB) to partition examples into subsets until nodes are pure or splitting yields minimal gain. During classification, examples are evaluated by traversing the nodes (i.e. queries), ending at leaves that assign a class (pos/neg) based on the majority label of examples in that leaf. The learned tree translates into a logical theory composed of rules of the form: `class([pos/neg]):-q1, ...,qn.` where the head is a single atom representing the majority class (positive or negative) of the leaf, while the body consists of the conjunction of queries (`q1,...,qn`) along the path from the root to that leaf. Aleph[5] is another ILP system that learns logic programs from positive $E^+$ and negative $E^-$ examples. Given a KB containing ground facts, Aleph builds hypotheses through a general-to-specific search strategy. It begins by selecting a seed example from $E^+$ and generating a saturation set, which represents an overly specific clause derived from the KB. This clause is then generalized via top-down refinement, iteratively adding literals to maximize coverage of $E^+$ while minimizing coverage of $E^-$, guided by a heuristic function. Each candidate clause is evaluated based on accuracy or coverage. If a clause meets predefined criteria, it is added to the hypothesis, and the covered positive examples are removed from

[4] https://dtai.cs.kuleuven.be/static/ACE/
[5] https://www.cs.ox.ac.uk/activities/programinduction/Aleph/aleph.html

$E^+$. The process repeats until all positive examples are covered or no further valid clauses can be found. The final theory is a disjunctive set of such clauses. For instance, a learned rule might take the form: `patient(X,pos):-b1,...,bn.`, where `b1,...,bn` are the literals allowed to appear in the body, while the target predicate is in the head.

PASCAL[6] is a PLP algorithm designed to perform both structure and parameter learning to discover Probabilistic Constraint Logic Theories (PCLTs) from labeled interpretations. A PCLT is a set of Probabilistic Integrity Constraints (PICs) of the form $p :: Body \rightarrow Head$, where $p \in [0, 1]$ represents the probability of the constraint. $Head$ is a disjunction of literals $L_1 \wedge ... \wedge L_n$, while $Body$ is a disjunction of positive or negative conjuncts $P_j$ and $N_j$: $\exists(P_1) \vee ... \vee \exists(P_n) \vee \forall\neg(N_1) \vee ... \vee \forall\neg(N_m)$, where each conjunct is a conjunction of literals. The approach for assigning a semantics to PCLTs is inspired by the distribution semantics [17]: a PCLT defines a distribution over non-probabilistic theories by assuming independence among the choices in probabilistic constructs. The distribution semantics is one of the most successful approaches in PLP; for further details see [15]. The probability is to be interpreted as the strength of the IC: a value $p_i$ means that the sum of the probabilities of the possible theories where a grounding of the constraint is present is $p_i$. Learning the parameters $p_i$ is done by gradient descent, while learning the structure by a beam search that generates and refines the ICs, starting from the most specific one and refining it based on a language bias (see below). Under this formalism, models assign a probability of being positive to new interpretations. This probability can be computed in a time that is logarithmic in the number of groundings of the constraints that are violated in an interpretation. In other words, PCLTs allow one to classify new patients as belonging to the positive class in a very efficient way.

TILDE, Aleph and PASCAL use a "language bias" to constrain the search and guide the generation of valid hypotheses.

### 4.2 Mining Association Rules

ARM methods work by identifying frequent patterns, which are then translated into interpretable rules. The FP-growth (Frequent Pattern-growth) [10] algorithm constructs a compressed FP-tree and recursively extracts frequent patterns using a divide-and-conquer strategy, eliminating the need for candidates' generation and repeated database scans, as in the well-known Apriori. Eclat [23] leverages a vertical data format, associating each item with its transaction ID list, and efficiently discovers patterns through list intersections, making it well-suited for sparse datasets.

## 5 Experiments

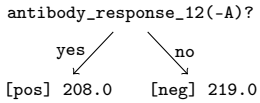
PASCAL experiments were run on the Leonardo HPC (Intel Xeon Platinum 8358, 2.60GHz), while TILDE and Aleph were executed on a local Linux VM

[6] https://friguzzi.github.io/pascal/_build/html/index.html

(AMD Ryzen 7 5800H, 3.19GHz, 6.2GB RAM) due to 32-bit requirements. For a fair comparison between learning times, TILDE and Aleph times were scaled by a factor of $\frac{3.19\text{GHz}}{2.60\text{GHz}}$ based on the clock frequencies. We used the subsample group of the study (611 patients), recording the immune response at timed intervals from the first dose. The dataset was partitioned into training and test sets following a 70-30% split, ensuring stratification to maintain a proportional representation of classes across both subsets. We also performed a 4-fold cross-validation; however, we observed that this did not lead to any noticeable improvement in performance or relevance of the theories. Two versions of the dataset were created: in the first one, interpretations were labeled as positive or negative according to criterion 1) mentioned at the end of Section 3.1; in the second one, we labeled interpretations based on criterion 2) (mortality). For TILDE and Aleph, we configured the language bias to match PASCAL's as closely as possible, despite some minor differences in the mode declaration syntax.

Table 1 shows the list of experiments relative to the (P)LP systems, specifying: the classification criterion, the number of positive and negative interpretations, the total number of intepretations, the number of training and test examples. It is important to note that, in 3 cases out of 4, the classes were imbalanced, which impacted the performances of the algorithms.

Performance evaluation took into account the Area Under the ROC Curve (AUC-ROC) [9], the Area Under the Precision-Recall Curve (AUC-PR) [6], the log-likelihood (LL) on training data and the learning time. Table 2 shows the results for the (P)LP systems. Aleph is not included as it failed to generate meaningful hypotheses, returning theories equivalent to the set of positive training examples and achieving an accuracy of 1. Even if TILDE achieves higher AUC-ROC and AUC-PR in tests 1–2, this is due to the trivial models returned, being able to classify most of the interpretations correctly. In fact, we obtained trees consisting only of the root node splitting into two leaves labeled pos and neg. For example, test #2 produced the following model:




PASCAL has very good or excellent performance across all experiments, particularly in cases with class imbalance (tests 3–4), where PCLTs perform perfect classification compared to TILDE's near-random guessing. On the other hand, the higher learning time of PASCAL is due to its much larger search space, as any predicate (not just the target) can be included in the head of the rules.

**Table 1.** Experiments' properties.

| Experiment | Classification | $I^+$ | $I^-$ | $I^{tot}$ | #Training Ex. | #Test Ex. |
|---|---|---|---|---|---|---|
| 1 | Antibody Response | 525 | 86 | 611 | 427 | 184 |
| 2 | Antibody Response | 275 | 336 | 611 | 427 | 184 |
| 3 | Death | 130 | 481 | 611 | 427 | 184 |
| 4 | Death | 130 | 481 | 611 | 427 | 184 |

**Table 2.** Performance comparison between PASCAL and TILDE in terms of Area Under the ROC Curve (AUC-ROC), Area Under the Precision-Recall Curve (AUC-PR), Log-likelihood (LL) and learning time in seconds. Bold values indicate the best results.

| | PASCAL | | | | TILDE | | | |
|---|---|---|---|---|---|---|---|---|
| Exp. | AUC-ROC | AUC-PR | LL | Time(s) | AUC-ROC | AUC-PR | LL | Time(s) |
| 1 | 0.835 | 0.961 | -57.4 | 1303.96 | **0.958** | **0.987** | **-55.7** | **0.15** |
| 2 | 0.823 | 0.749 | -104.6 | 916.91 | **0.961** | **0.912** | **-81.4** | **0.05** |
| 3 | **1** | **1** | **-4.9** | 197.00 | 0.500 | 0.211 | -319.2 | **0.61** |
| 4 | **1** | **1** | **-0.7** | 1264.94 | 0.600 | 0.351 | -225.4 | **3.06** |

FP-Growth and Eclat returned ARs based on metrics such as support, confidence, lift. Lift, in particular, was prioritized to identify non-trivial relationships: a value >1 indicates positive association, <1 negative association, and 1 independence between the antecedent and the consequent. We repeated learning by setting many different thresholds for support ([40–70%]) and confidence ([50–80%]), but we obtained only trivial rules.

In the following, the most relevant PICs learned by PASCAL are reported. The first 3 constraints come from test 1 while the 4th one is from test 2:

1) $0.163231 :: \text{delta\_antib\_resp26}(-94.75) \wedge \text{cognitive\_disorder} \wedge \text{delta\_antib\_resp612}(4457.0) \rightarrow$
$(\text{flu\_vaccine}(1) \wedge \text{hz\_vaccine}(0)) \vee \neg\text{who\_followup1}(0) \vee \neg\text{covid\_vaccine}(0, 2)$

2) $0.166981 :: \text{inflammatory\_bowel\_disease} \wedge \text{chronic\_kidney\_disease} \rightarrow$
$(\text{delta\_antibody\_resp02}(95.19) \wedge \text{delta\_antibody\_resp26}(-94.75))$
$\vee (\text{number\_doses}(2) \wedge \text{tot\_drugs}(8)) \vee (\text{delta\_anti\_resp26}(-94.75) \wedge \text{who\_f2}(0))$

3) $0.161470 :: \text{antibody\_resp\_6}(5.25) \wedge \text{thyroid\_disorder} \rightarrow$
$\neg\text{antibody\_resp\_12}(4490.0) \vee \neg\text{flu\_vaccine}(1) \vee \neg\text{covid\_vaccine}(0, 2)$

4) $0.370066 :: \text{outcome}(death) \rightarrow$
$\text{antibody\_resp\_6}(5.25) \vee \text{tot\_drugs}(8)$
$\vee (\text{antipneumovaccine}(0) \wedge \text{antibody\_resp\_6}(5.25))$

In PIC 1, cognitive disorder, a negative difference (around -94) in antibodies between T2 and T6 and a very high positive difference in antibodies between T6 and T12 imply the patient probably received the flu shot but not the herpes zoster one, or she showed a moderate to severe disease progression at the 1st follow-up T1 (who_followup1 $\neq 0$, where 0 means no infection), or she did not receive the second dose of the COVID-19 vaccine. In PIC 2, inflammatory bowel and chronic kidney diseases are probably linked to declining antibody responses over time (between T0-T2 (delta_antibody_resp02) and T2-T6 (delta_antibody_resp26)), or to polypharmacy and 2 doses of anti-COVID vaccine, or to no COVID infection at the 2nd follow-up (T2) and a reduction in antibodies between T2 and T6. This is supported by the scientific literature according to which, in older people, specific chronic diseases such as cognitive disorders or kidney failure can influence vaccine efficacy in terms of immune response and mortality risk. PIC 3 associates thyroid disorders and low antibody levels at 6 months with low antibody levels at 12 months or no flu shot or no 2nd dose of anti-COVID vaccine. In PIC 4, mortality risk is tied to low antibody

levels at 6 months or polypharmacy or no anti-pneumococcal vaccine. PIC 4 is associated with a higher probability than the previous ones, indicating that the constraint is stronger: according to the experts of the study, the fact that 2 out of the 3 identified predictors of death (low antibody levels at 6 months, anti-pneumococcal vaccine) are related to immunology shows that immunological factors play a key prognostic role in the older population. The dimension of uncertainty explored by PASCAL reflects heterogeneity, and heterogeneity is strictly age-related and peaks in the older population [13].

## 6 Conclusions

In this work we applied Probabilistic Logic Programming to learn models from missing and uncertain data collected in the GeroCovid Vax Study. Our method uncovers, both new and confirmed by literature, declarative probabilistic integrity constraints about factors affecting immune response in elderly populations post-COVID-19 vaccination, offering interpretable support to medical staff. Moreover, it is shown to outperform traditional non-probabilistic ILP and data mining techniques, being the only system to learn non-trivial knowledge.

## 7 Acknowledgements

Research funded by the Italian Ministry of University and Research through PNRR - M4C2 - Investimento 1.3 (Decreto Direttoriale n. 341 del 15/03/2022), Partenariato Esteso PE00000013 - "FAIR - Future Artificial Intelligence Research" - Spoke 8 "Pervasive AI" - CUP J33C22002830006, under the "NextGeneration EU programme". GeroCovid VAX study was sponsored by the Italian Medicines Agency (AIFA). The authors thank all the participants and health personnel who contributed to this study.

## References

1. Agrawal, R., Imieliński, T., Swami, A.: Mining association rules between sets of items in large databases. SIGMOD Rec. **22**(2), 207–216 (Jun 1993)
2. Altaf, W., Shahbaz, M., Guergachi, A.: Applications of association rule mining in health informatics: a survey. Artificial Intelligence Review **47**, 313–340 (2017)
3. Blockeel, H., De Raedt, L.: Top-down induction of first-order logical decision trees. Artificial Intelligence **101**(1), 285–297 (1998). https://doi.org/10.1016/S0004-3702(98)00034-4
4. Côrte-Real, J., Dutra, I., Rocha, R.: On applying probabilistic logic programming to breast cancer data. In: Lachiche, N., Vrain, C. (eds.) Inductive Logic Programming. pp. 31–45. Springer International Publishing, Cham (2018)
5. Côrte-Real, J., Mantadelis, T., Dutra, I., Roha, R., Burnside, E.: Skill - a stochastic inductive logic learner. In: 2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA). pp. 555–558 (2015). https://doi.org/10.1109/ICMLA.2015.159

6. Davis, J., Goadrich, M.: The relationship between precision-recall and roc curves. In: Proceedings of the 23rd International Conference on Machine Learning. p. 233–240. ICML '06, Association for Computing Machinery, New York, NY, USA (2006). https://doi.org/10.1145/1143844.1143874
7. Delgado, M., Sánchez, D., Martın-Bautista, M., Vila, M.: Mining association rules with improved semantics in medical databases. Artificial Intelligence in Medicine **21**(1-3), 241–245 (2001)
8. Doddi, S., Marathe, A., Ravi, S., Torney, D.: Discovery of association rules in medical data. Medical informatics and the Internet in medicine **26**, 25—33 (2001)
9. Fawcett, T.: An introduction to roc analysis. Pattern Recognition Letters **27**(8), 861–874 (2006). https://doi.org/https://doi.org/10.1016/j.patrec.2005.10.010
10. Han, J., Pei, J., Yin, Y., Mao, R.: Mining frequent patterns without candidate generation: A frequent-pattern tree approach. Data Mining and Knowledge Discovery **8**(1), 53–87 (Jan 2004)
11. Huang, J., Fox, J., Gordon, C., Jackson-Smale, A.: Symbolic decision support in medical care. Artificial intelligence in medicine **5 5**, 415–30 (1993)
12. Mandreoli, F., Ferrari, D., Guidetti, V., Motta, F., Missier, P.: Real-world data mining meets clinical practice: Research challenges and perspective. Frontiers in Big Data **5** (2022). https://doi.org/10.3389/fdata.2022.1021621
13. Nguyen, Q.D., Moodie, E.M., Forget, M.F., Desmarais, P., Keezer, M.R. et al.: Health heterogeneity in older adults: Exploration in the canadian longitudinal study on aging. Journal of the American Geriatrics Society **69**(3), 678–687 (2021)
14. Op De Beéck, T., Hommersom, A., Van Haaren, J., van der Heijden, M., Davis, J. et al.: Mining hierarchical pathology data using inductive logic programming. In: Artificial Intelligence in Medicine. pp. 76–85. Springer Int. Pub., Cham (2015)
15. Riguzzi, F., Bellodi, E., Zese, R., Alberti, M., Lamma, E.: Probabilistic inductive constraint logic. Machine Learning **110**, 1–32 (04 2021)
16. Sakhanenko, N.A., Galas, D.J.: Probabilistic logic methods and some applications to biology and medicine. Journal of Computational Biology **19**(3), 316–336 (2012)
17. Sato, T.: A statistical learning method for logic programs with distribution semantics. In: Proc. of the 12th Int. Conf. on Logic Programming (ICLP'95). pp. 715–729. MIT Press (1995)
18. Srinivasan, A.: The Aleph Manual (1999), http://www.cs.ox.ac.uk/activities/machinelearning/Aleph/aleph
19. Stilou, S., Bamidis, P., Maglaveras, N., Pappas, C.: Mining association rules from clinical databases: an intelligent diagnostic process in healthcare. Studies in health technology and informatics **84**(Pt 2), 1399—1403 (2001)
20. Trevisan, C., Del Signore, S., Fumagalli, S., Gareri, P., Malara, A. et al.: Assessing the impact of covid-19 on the health of geriatric patients: The european gerocovid observational study. European Journal of Internal Medicine **87**, 29–35 (2021). https://doi.org/10.1016/j.ejim.2021.01.017
21. Urbanowicz, R.J., Moore, J.H.: Exstracs 2.0: description and evaluation of a scalable learning classifier system. Evolutionary Intelligence **8**(2), 89–116 (2015)
22. Varghese, J., Beierle, C., Potyka, N., Kern-Isberner, G.: Using probabilistic logic and the principle of maximum entropy for the analysis of clinical brain tumor data. In: Proc. of the 26th IEEE Int. Symposium on Computer-Based Medical Systems. pp. 401–404 (2013). https://doi.org/10.1109/CBMS.2013.6627826
23. Zaki, M.: Scalable algorithms for association mining. ieee trans knowl data eng. Knowledge and Data Engineering, IEEE Transactions on **12**, 372 – 390