University of Applied Sciences and Arts Northwestern Switzerland
School of Business

member of swissuniversities

AACSB ACCREDITED

An Early Warning System that combines Machine Learning and a Rule-Based Approach for the Prediction of Cancer Patients' Unplanned Visits

H. F. Witschel*, E. Laurenzi*, S. Jüngling*, Y. Kadvany[θ] and A. Trojan[θ]

**FHNW University of Applied Sciences and Arts Northwestern Switzerland, Riggenbachstrasse 16, CH-4600 Olten*

[θ]mobile Health AG, Falkenstrasse 21, CH-8008 Zürich

27 March 2023, Symposium AAAI-MAKE, San Francisco (CA)

The Burden of Cancer Worldwide

- Cancer is among the leading causes of death worldwide.
- Generally, cancer rates are highest in countries whose populations have the highest life expectancy, education level, and standard of living.
- By 2040, the number of new cancer cases per year is expected to rise to 29.5 million and the number of cancer-related deaths to 16.4 million.

Cancer Tomorrow | IARC - - Data version: 2020

Source: International Agency for Research on Cancer

Problem

- Oncologists are under pressure to visit an increasing number of patients!
  - **Unplanned visits are problematic** (from a Swiss research project).
    - Hard to dedicate adequate time to complex cancer cases
      - Not optimal cancer patient outcomes

  - A possible approach: an early warning system to predict issues leading to unplanned visits.
  - A ML-based approach is proven to be successful for clinical risk prediction (Bull et al. 2020).
    - Problems (Ginestra et al. 2019):
      - reasons for raising alerts are not understood by clinicians AND
      - generic lack of trust of pure ML approach
  - Genuinely **human interpretable models can help** vs. relying on try to explain black boxes (Rudin 2019).

Our Approach and Underlying Assumptions

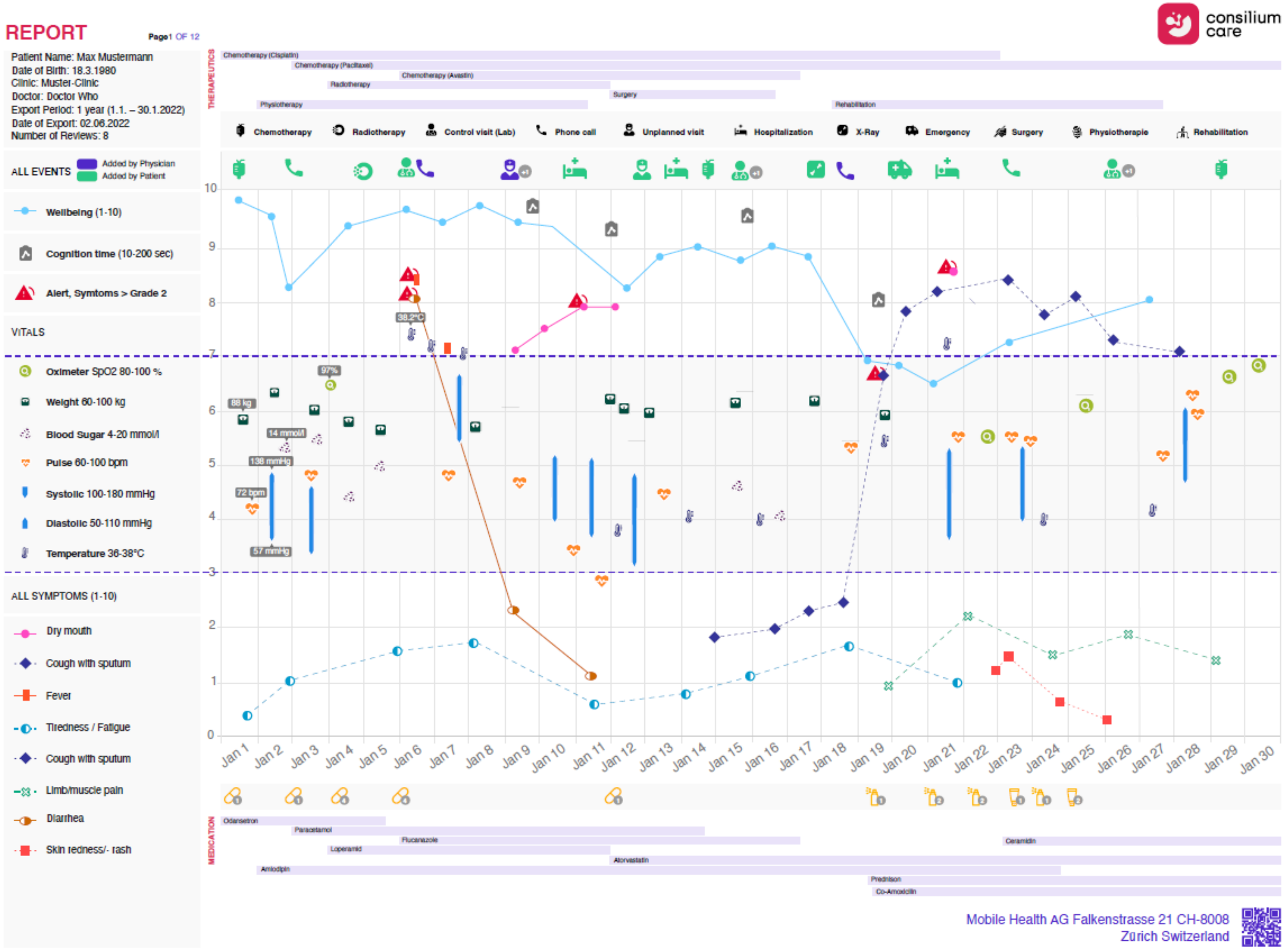
- An **interpretable Rule-based Machine Learning approach** that predicts issues leading to unplanned visits.
  - a learning algorithm that produces directly interpretable rules.
  - rules can be interpreted, formulated and enhanced by experts based solely on an understanding of input-output associations.

- **Assumptions:**
  - **Learned rules express conditions** for the doctors to advise a visit
  - **Doctors may not be able to externalize all rules**
  - **Doctors are able to judge** the **learned rules**
    - **checking and modifying ML-generated rules**
    - extend the rule base with low effort and at the same time use their **full medical knowledge** to control which cases will lead to warnings or not.

Data Set

– Diary kept by cancer patients in the form of a mobile app where data is entered on a daily basis.

- Total of 16’670 diary entries.
- 266 patients (mostly suffering by breast cancer).



Attributes used to train our EWS

- Ground truth: unplanned visits of patients
  - learn situations that lead to unplanned visits.
- 1 instance = **diary entry** = combination patient-day (i.e. training example)
- Association of class attribute **"unplanned visit"** to each instance
  - For each actual unplanned visit, we consider a **time frame of 3 days prior**, where the respective class attribute gets assigned the value "yes".
    - Time frame estimated by our oncologists.
  - **Drugs and symptoms strengths** are encoded
    - Patients received a guideline with def.
      - Developed by oncologists and
      - Based on a specific terminology (CTCAE)

Attribute(s)	number	description	type, value range
Birth year	1		numeric
Sex	1		{male,female}
Primary tumor	1		{breast, gut, blood/lymph, lung, prostate}
Wellbeing	1	subjective wellbeing of patient	numeric [0...100]
Therapy form	1	frequency of treatment	{daily, weekly, bi-weekly, 3-weekly, 4-weekly}
Drugs	88	cancer and other drugs	numeric [1,nan]
Symptom strengths	52	strength of relevant symptoms	numeric [0...1,nan]
Diagnosis terms	246	Terms occurring in patient's diagnosis	numeric (TF/IDF)
Note terms	311	Terms from patient notes	numeric (TF/IDF)
Unplanned visit	1	Class attribute	{yes,no}

- Free-text attributes
  - Diagnosis (entered by the doctor) AND Note (optionally entered by the patient)
  - String attributes were vectorised using TF/IDF weights -> prefixed with «diag_» and «note_», resp.

Proposed division of labor between ML and expert (1/3)

1. The human expert states a set of rules A.
   - workshop where the medical expert was invited to first look at a number of example cases where unplanned visits had happened. Next, the expert formulated some rules. E.g.,
     - Rule: diag_nikotine_abuse > 0) and (Chest_Pain > 0) => UnplannedVisit=yes
     - Problem category: Lung
2. We apply the rule learner to generate a set of rules B. E.g.,
   - Rule: (Wellbeing <= 46) and (Endoxan >= 1) and (diag_lymphangiosis >= 5.296978) => UnplannedVisit=yes
   - None of the learned rules matched with the rules suggested by the experts.
3. Rules from both A and B are evaluated based on a cost matrix where false negatives(FNs) have higher cost than false positives (FPs, false alarms). Rules are ranked by cost.

FN:
situations that should lead to unplanned visits are not recognized!

Rules	Total weight	ror weig	matches	incorrect (false positives)	correct	Precision	Recall	F measure	Cost
(diag_clipmarkierung >= 4.661854) and (Wellbeing <= 74) => UnplannedVisit=yes	147.76	37.63	53	41	12	22.6%	7.2%	0.110	1581
(Wellbeing <= 46) and (Endoxan >= 1) and (diag_lymphangiosis >= 5.296978) => UnplannedVisit=yes	71.58	7.34	15	8	7	46.7%	4.2%	0.077	1598
(Schmerzen >= 0.38) and (note_heute >= 2.30896) and (Wellbeing <= 60) => UnplannedVisit=yes	73.42	9.18	17	10	7	41.2%	4.2%	0.077	1600
(Schmerzen >= 0.37) and (Appetitverlust <= 0.43) and (Müdigkeit >= 0.44) => UnplannedVisit=yes.	62.41	7.34	14	8	6	42.9%	3.6%	0.067	1608
(Wellbeing <= 41) and (Schlafstörung <= 0.61) and (Atemnot <= 0.58) and (diag_1 >= 1.267966) => Unplanned	48.64	2.75	8	3	5	62.5%	3.0%	0.057	1613
(BirthYear <= 1954) and (Wellbeing <= 65) and (Fieber <= 0.18) => UnplannedVisit=yes	71.58	16.52	24	18	6	25.0%	3.6%	0.063	1618

Performance of machine-learned rule set

- Our machine-learned rule set (cost ratio 1:10):
  - discovered **47** out of the 166 critical situations (28.3% recall)
  - generated **263 false alarms** (15.2% precision).
- When we use a **ratio of 1:20** for the cost-sensitive rule learner:
  - the model **recognised 54 critical situations,** i.e. 7 more than with the 1:10 model,
  - **but at a cost of an extra 104 false alarms**, i.e. overall 367 false positives

		Predicted visit	
		Yes	No
Unplanned visit	Yes	47	119
	No	263	16241

Confusion matrix (1:10)

Proposed division of labor between ML and expert (2/3)

4. ML rules are inspected by the human expert in the ranked order. The expert can suggest to drop a rule, but also to modify it, e.g., dropping or adding a condition. Modified rules will be evaluated and accepted if their cost on the test set is acceptable.

- Out of the **12 rules**, 3 were accepted in their original version.
- Another **3 rules were rejected** - symptoms were not critical or unclear interpretation.
- The **remaining 6 rules were accepted with modifications**
  - **In 2 cases**, these modifications were **additional conditions (**e.g., additional symptoms that would make a situation truly critical). The additional conditions involved new attributes related to the trend, i.e.,
    - it will require to construct new features that take the **historical development of symptom strengths** into account.
  - **In the remaining 4 cases**, we discovered new insights:
    - predict a **situation that requires action (a different kind of alert)**, but not necessarily a visit. Not possible when working with black-box machine learning models!!!

Example of two learned and interpreted rules

- **Rule 1**: (diag_clipmarker >= 4.661854) and (Wellbeing <= 74) => UnplannedVisit=yes
  - **Medical interpretation:** that breast cancer patients receive a neoadjavant chemotherapy before a surgery, during which the tumor shrinks (which is why its position is marked with a "clip marker" – i.e., that term correlates with a specific treatment). The chemotherapy impacts the wellbeing negatively.
  - **Action:** suggest a visit if two additional conditions apply:
    - a) the trend of wellbeing is negative over several days and,
    - b) when nausea or fatigue appear as accompanying symptoms.

*example of human recommended **additional conditions** in ML-discovered rules*

- **Rule 2:** (Wellbeing <= 46) and (Endoxan >= 1) and (diag_lymphangiosis >= 5.296978) => UnplannedVisit=yes
- **Medical interpretation:** situation of palliative care
- **Suggested actions:** instead of a visit, advise the home care service to intensify measures to alleviate the suffering and ensure a higher wellbeing. E.g., increase the dose of painkiller medication or reduce that of Endoxan

*example of human recommended **new type of alert** in ML-discovered rules*

Conclusion

- We have proven that a machine learning approach to the discovery of rules may provide value and that a corresponding model will be able to discover several critical situations.

- The utility of a machine-learned rule set will be limited because increasing its coverage (recall) is possible, but comes at the cost of lower precision, i.e. more false alarms.

- The analysis of rules by medical experts not only results in rule modifications and suggestions for feature engineering, but also in **specific types of actions** entailed by predictions that allow for more fine-grained recommendations to be made by the Early-Warning System.

Outlook

- Advanced feature engineering
  - Consider the history of well-being, symptom development over time, missing patient entries of previous days
- Knowledge engineering
  - Grouping symptoms, drugs and their connection meaningfully
    - including the resulting symptom categories as new features.

Thank you.

emanuele.laurenzi@fhnw.ch