UCLouvain
Institute of Information and Communication Technologies, Electronics and Applied Mathematics

icteam

Bringing Highly Specialized Medical Expertise in General Practitioners' EHR to Reduce Patients' Diagnostic Time for Rare Diseases

François Roucoux *, Sébastien Jodogne *
* ICTEAM, INGI (UCLouvain)

Context

Rare Diseases, a Critical Health Issue

Individually rare (< 1 case / 2000 patients) but quite common as a whole [1]

1 out of 12 People affected in Europe

6200+ clinically described

80% of genetic origin

70% manifest in childhood

4-8 years until a correct diagnosis

Ultimately Fatal

Mostly chronic, degenerative, disabling

GPs are Essential for their Screening

They constitute the first point of contact of the population with the health system

They collect information and biomarkers and store them in **EHR**

They refine differential diagnoses

They detect red flags

They request additional diagnostic tests

They refer patients to rare disease specialists

They coordinate long term follow-up

But, **GPs lack expertise in rare diseases screening [2]**

Objectives

Provide rare diseases (RD) screening expertise to GPs to:

- **reduce the time** taken for the RD diagnosis
- **improve the quality** of care for RD-affected patients
- **limit the progression** of the RD before treatment

Focus on the **four rare diseases** which are the subject of an agreement with the Belgian National Health Insurance:

- **Epidermolysis bullosa**: inherited skin fragility
- **Primary immunodeficiency**: alteration of immune system components
- **Multiple system atrophy**: adult disease of the central nervous system
- **Idiopathic pulmonary fibrosis**: lung damage in older adults

(+ other rare diseases given medical collaboration opportunities)

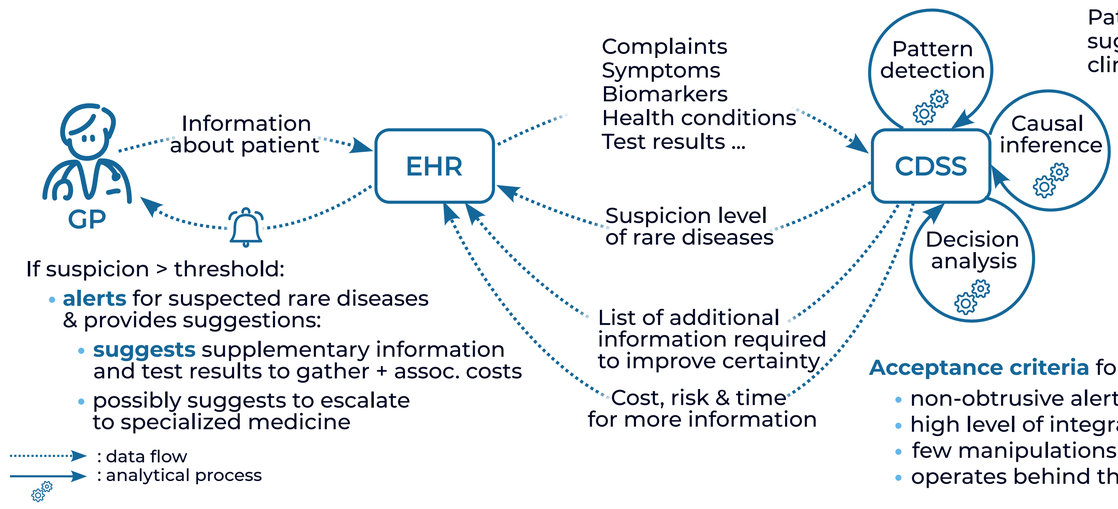
Take into account the **health economics dimension** into the screening model (suggested tests, genetic analyzes)

Perform a **clinical study** to prove reduction of diagnostic time and associated positive outcomes for patients

[1] The Lancet Global Health Redaction. The landscape for rare diseases in 2024. *The Lancet Global Health*, 12(3):e341, 2024.
[2] W. R. Evans and I. Rafi. Rare diseases in general practice: recognising the zebras among the horses. *British Journal of General Practice*, 66(652):550–551, 11 2016.
[3] Daphne Koller and Nir Friedman. *Probabilistic graphical models: principles and techniques*. MIT press, 2009.
[4] Orphadata. https://www.orphadata.com/. Accessed: 2024-06-29.
[5] Mark L Braunstein. *Smart on fhir. In Health Informatics on FHIR: How HL7's API is Transforming Healthcare*, pages 293–325. Springer, 2022.

Methods

Coupling a Clinical Decision Support System (CDSS) with GP's EHR



Pattern detection spots patterns suggestive of rare diseases in clinical observations in terms of:

- Occurence
- Recurrence
- Temporality

Acceptance criteria for the CDSS:

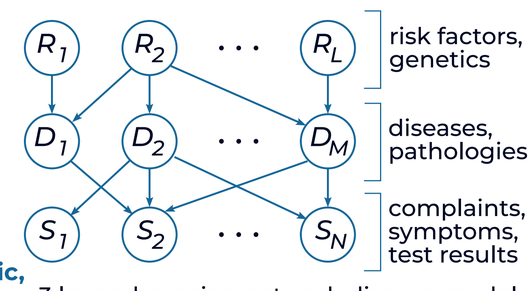
- non-obtrusive alerting system
- high level of integration with the EHR data
- few manipulations required on the part of the GP
- operates behind the scene of the EHR

Bayesian Networks to Model the Domain Knowledge of Rare Diseases

Bayesian Networks (BN) are used to model domain knowledge as a **directed acyclic graph** (DAG) representing **probabilistic relationships** among a set of **variables** [3].

Directed edges represent **causal influences** among its variables (nodes).

Widely used in Medicine for: **diagostic, pronostic, decision making** for therapy...



3 layers bayesian network disease model

Modeling challenge

Require **a lot of numerical parameters** to populate their conditional probability tables (CPT) quantifying causal relationships.

A binary variable with n binary predecessors requires 2^n independent parameters.

e.g. : S_2, influenced by D_1, D_2 and D_M require 2^3=8 parameters to determine its value.

Bayesian Noisy-OR Gate for Causal Inference

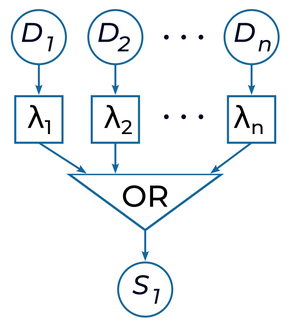
Reformulate each layers of the BN domain model into a Noisy-OR Gate.

It provides a **logarithmic reduction** [3] in the number of parameters required.

It works given the following **assumptions**:

- Limit variables to **binary** ones.
- **Causal independence of causes**: each cause contributes to the effect in an independent manner, such that the influence of one cause does not depend on the presence or absence of other causes.

Causal independence may be difficult to guarantee, particularly for risk factors



Noisy OR-gate model for diseases and symptom

The probability that symptom *S* occurs (i.e., S=1) is 1 minus the product of the failure probabilities (λ) of the active diseases $D_{1..n}$:

$$P(S = 1 \mid D_1, D_2, \ldots, D_n) = 1 - \prod_{D_i=1} \lambda(D_i)$$

λ computes the chance that the cause D_i alone fail to produce the effect

Calculate the posterior probability of the cause (disease) given an observed effect (symptom) involves applying Bayes' Theorem to invert the direction of the inference:

$$P(D_j = 1 \mid S = 1) = \frac{P(S = 1 \mid D_j = 1) \cdot P(D_j = 1)}{P(S = 1)}$$

(Formula extendable to multiple diseases)

Influence Diagrams as Decision Model

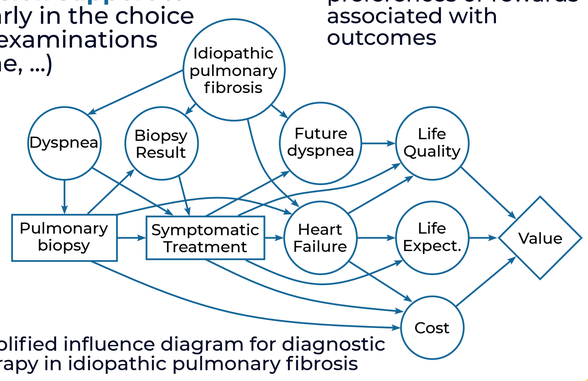
Extends the concept of Bayesian network by incorporating not only probabilistic relationships between variables but also **decisions** and **objectives** (or utilities) [3].

Provides **decision support** to GPs, particularly in the choice of additional examinations (cost, risk, time, ...)

Chance node: random variables or uncertain events

Decision node: choices or actions available

Utility node: objectives, preferences or rewards associated with outcomes



E.g. simplified influence diagram for diagnostic and therapy in idiopathic pulmonary fibrosis

Standard-based Integration

orphanet **World reference knowledge base** about rare diseases.

To feed inference and decision models of the CDSS either through its **standardized API** [4] or through an intermediate **knowledge modeling** step.

Standard-based integration between EHR & CDSS:

- **Semantic interoperability**[5] ensuring that concepts are understood consistently between EHR and CDSS

SNOMED CT ICD LOINC

- **Technical interoperability**[5] ensuring consistent data formats and protocols between EHR and CDSS

data format & exchange protocol

Summary

- Rare diseases (RD) constitute an important health problem.
- General practitioners (GPs) can play a determining role in their screening.
- GPs lack the expertise to carry out this screening, increasing time before diagnosis.
- Coupling a clinical decision support system (CDSS) with their EHR shall provide them part of this expertise.
- A Bayesian model can be feeded with statistical knowledge (risk, symptoms, incidences ...) from Orphanet.
- Bayesian noisy OR-gate model of diseases allows causal inference without requiring to much parameters.
- Influence diagrams shall provide decision support (risk, cost ...) about supplemental tests or required actions.
- Technical and semantic interoperability standards ensure sound integration between EHR and CDSS.