|  |  |  |
| --- | --- | --- |
| ITU Logo | INTERNATIONAL TELECOMMUNICATION UNION**TELECOMMUNICATIONSTANDARDIZATION SECTOR**STUDY PERIOD 2017-2020 | FG-AI4H-H-028 |
| **ITU-T Focus Group on AI for Health** |
| **Original: English** |
| **WG(s):** | Plenary | Brasilia, 22-24 January 2020 |
| **DOCUMENT** |
| **Source:** | Anastasia.ai & Tecnigen |
| **Title:** | Proposal for new topic group: Deep learning model profiling and risk score for diabetes mellitus type 2 and pre-diabetes and their complications |
| **Purpose:** | Discussion |
| **Contact:** | Dr. Andrés ValdiviesoAnastasia.ai & TecnigenInnovation Department in Anastasia.aiSantiago, Chile | Tel: +56996330961Fax: +Email: avaldivieso@anastasia.ai  |

|  |  |
| --- | --- |
| **Abstract:** | This document proposes a new topic group on a deep learning model in diabetes mellitus and pre-diabetes diagnostics and DCSI score range with a focus on the identification and distribution of scoring of it complication for kidney and heart problem improving the scales and decision trees that are used today for diabetes and pre-diabetes in real-time and in an aggregate way for population health. |

**Overview**

The Diabetes Complication Severity Index (DCSI) is used from clinical laboratory data and ICD-10 to predict complications, mortality and hospitalizations of diabetic patients, this system is used by doctors through a system of patient risk score scores to later group them according to their risk and take action behaviors

The model has certain limitations

1) Visibility of Medical to Medical Information and not in aggregate form which limits immediate population health decision making, apart from a case-by-case interpretation

2) Loss of visibility of patient risks between distributions in the risk scale (example between 0 and 1 and 1 and 2) since the score only shows integer intervals.

3) Decision tree system, these systems are based on a decision tree of the scores without including other exams or factors that could modify the score and behaviors

Although this score is generally evaluated by manual estimation of physicians, efforts are being made to evaluate DSCI and its recommendations through automatic deep learning analysis methods from the online information of laboratory equipment. These methods, and particularly those based on AI, are still experimental and not sufficiently documented and standardized for introduction into clinical trials and in daily practice and are being used for guidance.

Therefore, we propose to establish a set of data for the comparative evaluation of DSCI detection algorithms of diabetic and pre-diabetic patients, their risks and recommendations to be followed based on the above using automatic learning and quantification of DCSI for chronic patients

**Impact**

Chronic diseases, such as diabetes mellitus (DM), hypertension and cardiovascular diseases, with a prevalence nationwide, diabetes is 12.32%, HTA prevalence it is 26.9% (source national health survey)[[1]](#endnote-1)

It is estimated that the costs associated with the treatments and complications of these diseases correspond to 82% of each dollar spent on health.[[2]](#endnote-2)

In a study, the average U.S. healthcare spending per person in 2014[[3]](#endnote-3), there is 30% of wasted spending due to unnecessary services, inefficiencies, and fraud, among others.

“Clinical practice guidelines have been developed for many common conditions based on data from randomized controlled trials. When medicine is informed solely by clinical practice guidelines, however, the patient is not treated as an individual, but rather a member of a group.” [[4]](#endnote-4)

Healthcare is facing a scenario in which there are patients treated as average individuals, with average diagnostics, treatments, and results, to whom are being defined services that result in inefficient use of resources and efforts for both the institution, but most of all, and the patient.

The need for personalized has increased over the years and, as in the retail and consumer markets, patients have become more expert in health/costs and question their health status, treatments, and associated costs.

In the health industries, institutions are generating and storing so much data, but they have not fully exploited or extracted the value of a deeper analysis of the data to understand and treat each patient as a unique individual.

However, this large amount of data cannot be stored by the doctor's memory and will have to rely on solutions that will improve their abilities to be effective and efficient to do a better job in the future. And on the other hand, seeks to change from curative medicine to a preventive and to be able to anticipate the complications of chronic diseases.

The tables below show the possible cost impact of preventing complications due to better actions by having more segmented knowledge of patients during their total life cycle.



The magazine shows the differences in the subgroups of type 2 “Identification of type 2 diabetes subgroups through topological analysis of patient similarity from *Science Translational Medicine* 28 Oct 2015: Vol. 7, Issue 311, pp. 311ra174” so try to cluster and achieve a greater number of subsegments to improve characterization and personalized actions on them.

Currently, the problem is solved by the population health challenges through control panels that measure the risk criteria of some chronic diseases, the disadvantage of this is that the values are absolute for them does not consider the differences between for example a patient who has 200 Cholesterol v / s one who has 195, which leads to the analysis at wider bands, the second does not associate in real-time the history of events and tests that may be associated and that if they are not informed by the patient Or the doctor does not possess enough experience they will not be perceived. This information is not always up to date and/or a small group of people have access to its visualization and searching for each patient is not so simple, so that it is consulted by primary care doctors or managers without knowledge of analytics. One of the barriers to this dashboards is the dispersion of the data and its reliability since many of them are not complete since they try to be extracted from the clinical records (and not all of them are complete) or they are not directly owned in hospitals tertiary or primary health center.

**Data Availability**

Currently, no high-quality annotated data sets on diabetes and pre-diabetes score distribution of complications (DSCI) are publicly available.

But the data source for this model is available in any hospital clinical laboratory system (LIS) initially being supplemented in a second stage by EMR

The interesting thing about the initial data sets from the LIS is that the entropy they present is the highest in order to generate a deep learning training compared to other sources of health data sources

We intend to provide a complete set of DSCI score data that allows the evaluation of data analysis methods for the calculation of DSCI for diabetes and pre-diabetes. These data will be provided anonymously within a computer infrastructure in HIPPA standard which will be used for the actual benchmarking process.

We will provide a second set of data (anonymous and smaller) for public download for participants to evaluate the general characteristics of the data used for benchmarking.

**Benchmarking**

In the benchmarking process, participants are expected to present AI-based solutions to analyze laboratory system data for

1. AI-powered Lab Operational Management to understand and take actions over right/wrong/overused/underused prescriptions

2. AI Pathology Management Dashboard for an overall distribution and score of DSCI view of population behavior around the pathologies trained to the solution (in this case, for Diabetes Pre Diabetes, Heart Conditions,)

3. Advanced AI Medical Assistant which helps the physician understand the specific characterization of his/her patients and the requirements to deliver personalized treatment. With this tool the doctor will have a summary of his patient history, personal risk KPIs and the explanation about those and, finally, recommended analysis to be made for each one.

Presentations should be evaluated by comparing predictions based on AI with manual annotations and scores given by physicians.

**Organizer Details**

The Joint Venture of Anastasia (company with more than 23 years in artificial intelligence work), Tecnigen laboratory system company founded in 1940 and the conception hospital the formal birth of an enclosure called Regional Hospital of Concepción goes back to 1943. There is built a five-story building and an underground, which was designed for 600 hospitalization beds and a small polyclinic for the attention of the time, all sheltered in more than 20 thousand square meters.

Due to the increasing requirements to provide more timely, quantitative and standardized and personalized diagnoses, our main interest is to promote and support the efforts of artificial intelligence techniques in order to help deliver tools that help to better manage and control chronic diseases such as diabetes and manage to arrive in advance to control pre-diabetes and manage its possible future complications worldwide

**References**

<https://towardsdatascience.com/machine-learning-for-diabetes-562dd7df4d42>

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5257026/>

<https://www.kaggle.com/uciml/pima-indians-diabetes-database>

<https://hackernoon.com/ml-for-diabetes-from-bangladesh-d99d1d058d82>

<https://www.fundaciondiabetes.org/general/noticia/14091/primer-proyecto-que-usa-la-inteligencia-artificial-para-tratar-la-diabetes>

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3810070/>

<https://arxiv.org/abs/1801.04062>

<https://arxiv.org/abs/1808.06670>

Libros y Articulos

The diabetes challenge in Chile Developing Knowledge-enhanced Chronic Disease Risk Prediction Models from Regional EHR Repositories (Jing-Mei, Ph.D.)

Machine learning for chronic disease (Katherine E. Niehaus and David A. Clifton)

PEPPER: Patient Empowerment Through Predictive Personalised Decision Support (Pau Herrero, Beatriz Lopez y Clare Martin)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. [www.idf.org/diabetesatlas](http://www.idf.org/diabetesatlas) [↑](#endnote-ref-1)
2. Medical Expenditures Panel survey ahrq 2012 [↑](#endnote-ref-2)
3. Consumer Reports - [Outrageous Health Costs](https://advocacy.consumerreports.org/wp-content/uploads/2014/09/Infographic5-1024x791.jpg) [↑](#endnote-ref-3)
4. Nacional Center for Biotechnology Information - [Personomics: The Missing Link in the Evolution from Precision Medicine to Personalized Medicine](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5748623/) [↑](#endnote-ref-4)