Artificial intelligence and Big Data in Diabetes care: A Position Statement of the Italian Association of Medical Diabetologists (AMD)

Nicoletta Musacchio, Annalisa Giancaterini, Giacomo Guaita, Alessandro Ozzello, Maria A Pellegrini, Paola Ponzani, Giuseppina T Russo, Rita Zilich, Alberto de Micheli

Submitted to: Journal of Medical Internet Research on: November 05, 2019

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on it's website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressively prohibit redistribution of this draft paper other than for review purposes.
Table of Contents

Original Manuscript............................................................................................................................................................................... 4
Artificial intelligence and Big Data in Diabetes care: A Position Statement of the Italian Association of Medical Diabetologists (AMD)

Nicoletta MusacchioMD, ; Annalisa GiancateriniMD, ; Giacomo GuaitaMD, ; Alessandro OzzelloMD, ; Maria A PellegriniMD, ; Paola PonzaniMD, ; Giuseppina T RussoMD, ; Rita ZilichMD, ; Alberto de MicheliMD,

Corresponding Author:
Nicoletta MusacchioMD,
Phone: +3902 5799 ext 4304
Fax: 39 02 57994309
Email: nicoletta.musacchio@gmail.com

Abstract

Since the last decade most of our daily activities have turned out to be digital. Digital health takes into account the ever increasing synergy between advanced medical technologies, innovation and digital communication. Thanks to Machine Learning we are not limited anymore to a descriptive analysis of the data, as we can obtain greater value through the identification and prediction of patterns resulting from inductive reasoning. Machine Learning software that disclose the reasoning behind a prediction allow for “what-if” models by which it is possible to understand if and how, by changing certain factors, one may improve the outcomes, thereby, identifying the optimal behavior.

Today diabetes care is facing several challenges: the decreasing number of diabetologists, the increasing number of patients, the reduced time allowed for medical visits, the growing complexity of the disease, both from the clinical and the patient-care standpoint, the difficulty of achieving the relevant clinical targets, the growing burden of disease management for both the healthcare provider and the patient, the healthcare accessibility and sustainability. In this context, the new digital technologies and the use of the artificial intelligence, certainly are a great opportunity.

Herein we report the result of a careful analysis of the current literature and represents the vision of the Italian Association of Medical Diabetologists (AMD) on this controversial topic that, if well used, may be the key for a great scientific innovation. AMD believes that the use of artificial intelligence will allow to turn data (descriptive) into knowledge of the factors that “affect” the behavior and correlations (predictive), thereby identifying the key aspects that may establish an improvement of the expected results (prescriptive). Artificial Intelligence can therefore become a tool of great technical support to help diabetologists, to become fully responsible of the individual patient, assuring customized and precise medicine. This, in turn, will allow for comprehensive therapies built in accordance with the evidence criteria that should always ground any therapeutic choice.

(JMIR Preprints 05/11/2019:16922)
DOI: https://doi.org/10.2196/preprints.16922

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✔ Please make my preprint PDF available to anyone at any time (recommended).
  Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
  Only make the preprint title and abstract visible.
  No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✔ Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).
  Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain visible.
  Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in
  the JMIR PubMed Now! service, my accepted manuscript PDF will automatically be made openly available.
Original Manuscript
Artificial intelligence and Big Data in Diabetes care: A Position Statement of the Italian Association of Medical Diabetologists (AMD)

N. Musacchio¹#, A. Giancaterini², G. Guaita³, A. Ozzello⁴, M.A. Pellegrini⁵, P. Ponzani⁶, G.T. Russo⁷, R. Zilich⁸, A. De Micheli⁹

Document approved and validated by the National Executive board of AMD 2015-2017*

1 Past President of the Associazione Medici Diabetologi (AMD), President of “Fondazione AMD” - Italy
2 Diabetology service, Muggiò Polyambulatory, Azienda Socio Sanitaria Monza, Monza - Italy
3 Diabetology, endocrinology and metabolic diseases service, ATS Sardegna-ASSL Carbonia - Italy
4 SSVD Diabetologia e Malattie Metaboliche, Ospedali Riuniti di Pinerolo ASL T03, Pinerolo - Italy
5 CdA Fondazione AMD, New Coram Udine, Udine – Italy
6 Operative Unit of Diabetology, La Colletta Hospital, ASL 3, Genova - Italy
7 Department of Clinical and Experimental Medicine, University of Messina, Messina – Italy
8 Partner Mix-x, Milano - Italy
9 ACISMOM, Genova - Italy

* National Executive board AMD 2015-2017: N. Musacchio, D. Mannino, A. Botta, M. Calabrese, F. Calcaterra, R. Candido, A. Gigante, G. Guaita, G. La Penna, S. Leotta, A. Lo Presti, M.C. Ponziani, M. Ragonese.

Running head: Artificial intelligence and Big Data in Diabetes care

# Corresponding author: Nicoletta Musacchio, MD;

e-mail: nicoletta.musacchio@gmail.com
Phone: (+39) 02 57994304/4306
Fax: (+39) 02 57994309
Abstract

Since the last decade most of our daily activities have turned out to be digital. Digital health takes into account the ever increasing synergy between advanced medical technologies, innovation and digital communication. Thanks to Machine Learning we are not limited anymore to a descriptive analysis of the data, as we can obtain greater value through the identification and prediction of patterns resulting from inductive reasoning. Machine Learning software that disclose the reasoning behind a prediction allow for “what-if” models by which it is possible to understand if and how, by changing certain factors, one may improve the outcomes, thereby, identifying the optimal behavior.

Today diabetes care is facing several challenges: the decreasing number of diabetologists, the increasing number of patients, the reduced time allowed for medical visits, the growing complexity of the disease, both from the clinical and the patient-care standpoint, the difficulty of achieving the relevant clinical targets, the growing burden of disease management for both the healthcare provider and the patient, the healthcare accessibility and sustainability. In this context, the new digital technologies and the use of the artificial intelligence, certainly are a great opportunity. Herein we report the result of a careful analysis of the current literature and represents the vision of the Italian Association of Medical Diabetologists (AMD) on this controversial topic that, if well used, may be the key for a great scientific innovation. AMD believes that the use of artificial intelligence will allow to turn data (descriptive) into knowledge of the factors that “affect” the behavior and correlations (predictive), thereby identifying the key aspects that may establish an improvement of the expected results (prescriptive). Artificial Intelligence can therefore become a tool of great technical support to help diabetologists, to become fully responsible of the individual patient, assuring customized and precise medicine. This, in turn, will allow for comprehensive therapies built in accordance with the evidence criteria that should always ground any therapeutic choice.

Keywords: artificial intelligence; big data analytics; clinical decision making; diabetes management; healthcare
Introduction

The ongoing evolution in medicine and in particular in the field of diabetology is strongly intertwined to a series of changes and innovations. The term *Digital Health* is a “container” grouping together informatics and telecommunications, that have the common objectives of diagnosis, treatment or monitoring of diseases, maintenance of health and well-being and support for healthy lifestyles. The U.S. Food and Drug Administration has compiled a list of software with medical device function (applications for smartphones and personal computers with diagnostic, monitoring or therapeutic objectives), advanced business intelligence data analysis tools, *artificial Intelligence (AI)*, cloud, cyber security, innovative health technologies [1]. These are tools belonging to an evolving reality, still unclear, potentially leading to challenging and promising scenarios [2].

The scope of this position statement is to analyze the most relevant of these aspects and describe the changes that have already occurred or those that will take place shortly, exploring the possibilities of application and development in the field of diabetology. Indeed, the digital world is constantly expanding and it has already become an integral part of our personal and professional life: smartphones, personal computers and network access are now essential tools for almost all the population and for many diabetologists, both as individuals and professionals.

Digitalization

Almost every daily activity of the diabetologist has to deal with *digitalization*, from electronic medical records to imaging diagnostics, from laboratory references to the various software for administrative practices and certifications, from glucose data download from glucometers to sensors for continuous glycemic monitoring and insulin pumps. Through the analysis of these data we already take therapeutic decisions and the more complete and numerous the data, the more we need informatics tools to guide us in the analysis, help us identify specific patterns for detecting glycemic abnormalities, understand the possible causes and adopt the appropriate therapeutic strategies to correct them.

Data management and connectivity

In the early days, *data collection and management* occurred only in the clinic, by cable or wireless devices; nowadays, thanks to technological evolution, these data can be transmitted automatically from the glucometers via “cloud” or on data integration platforms that collect elements from different devices and provide standardized reports. If not properly managed, this large amount of data is likely to overwhelm both the patient and the healthcare professional, and, indeed, the professional's experience in reading these reports is certainly the added value that increases their potential, and advanced informatics tools can simplify the analysis and provide suggestions to guide the clinical decisions of the physician. With experience and technology coming together, analysis time can be increasingly reduced, and more appropriate and correct data-driven therapeutic decisions are made, thanks also to greater customization strategies of the therapy.

Artificial intelligence and individualized care

In this context, the use of *AI*, and in particular of *Machine Learning*, allows an important step forward compared with more traditional data analysis techniques (i.e. graphs representing a picture of reality, very useful and precise, but static and outdated). Through the automatic identification of specific patterns within the data, and through the inductive reasoning typical of the human mind, *Machine Learning* can highlight correlations that lead to “predictions”, without being programmed in advance to carry out this activity.

In a not too distant future *AI*, thanks to algorithms able to learn and improve their own abilities
independently, will offer effective solutions to satisfy the most disparate needs and will come to deal with problems that today may seem insurmountable obstacles, for the benefit of the community [3]. In the field of clinical diabetology, these tools could have multiple potential implications, including the identification of new risk factors for the onset of diabetes (through the evaluation of large databases related to the general population), or revealing unsuspected subjects at high risk of complications (by cross-checking clinical and administrative records of diabetic patients), and identifying the behavioral and therapeutic variables most closely related to the progression of a specific complication.

There are many examples of collaboration between pharmaceutical companies, IT companies, scientific institutes and universities that exploit large complex datasets (the so-called Big Data) with the aim of improving the treatment of Type 2 Diabetes Mellitus and to unravel physio-pathological mechanisms, through the integration of data from biological, demographic, clinical, environmental and genomic sources.

**Risk stratification and personalized medicine**

The ability to elaborate, even in real time, through increasingly powerful algorithms, a large and heterogeneous amount of data allows us to extract knowledge and, exponentially, to make predictive assessments on the behavior of individuals as well as, more generally, to assume decisions for the entire community. This means that the use of AI will make it possible to transform the immediate data (descriptive), into knowledge of the factors that “condition” behaviour and correlations (predictive) up to the identification of the key factors that can facilitate an improvement in the expected results (prescriptive).

In this scenario, one of the most interesting application would be to identify which variables could be related to greater responsiveness to a specific drug, because this approach would open the door to a truly personalized medicine which uses the right drug for the right person, with obvious greater efficacy, improved outcomes and containment of costs. Furthermore, predictive analysis techniques based on AI could be used to identify which group of patients would require more attention and which strategies would be more effective depending on the individual patient, allowing more efficient methods of personal care, with lower costs and better outcomes. This concept of risk stratification, that considers all the individual characteristics of the patient (clinical aspects, genetic data, lifestyle, environmental factors), is the basis of a modern model of clinical governance called “population health management”. Today, the risk stratification of the population results from the extraction of historical expenditure data (in- and out-patients’ admissions, diagnosis, pharmaceutical expenditure) and this presents a series of biases: the clinical validity is weak, there is a temporal misalignment between the extraction, the analysis and the health status, information on the real socio-economic condition and behavioral health is absent. In this way, the data represent “satisfied demand” rather than “real needs”. The new challenge is the use of Big Data and business analytics systems to activate risk stratification models based on real health status, integrated use of multiple sources and collaboration between professionals (care team and data manager, in a data mining process).

**Biotechnologies and “omics”**

Another aspect that characterizes the ongoing medical revolution is the development and increasing importance of biotechnologies used in genetics and in the so-called “omics” sciences. These new techniques have been showing that the complex pathophysiological processes underlying Type 1 Diabetes mellitus (T1DM), T2DM and Gestational Diabetes Mellitus are caused by disturbances in gene expression that lead to alterations of processes within the organs involved in glucose homeostasis. The complexity of the system is exacerbated by the fact that the relative contribution of each component is highly individual. Understanding the molecular mechanisms underlying these
interactions is crucial for developing new personalized prevention and treatment strategies [4,5].

**Social Networks and Apps**

Finally, the ever-increasing use of *social networks* and *apps* has major implications in the diabetological reality. On one hand, health professionals must change the way they communicate and keep up with the times, always acquiring new technological and communication skills to deal with and manage the changes taking place, on the other, the new technologies could be used as a motivational and coaching tool, in support of traditional educational activities, and as an alternative data source. Notably, the areas that influence the state of health are only 10% due to medical care and health status, strictly; the major role is instead played by lifestyle and behaviors, as well as by genetics. For this reason, to get a more precise picture of the state of health of the general population, or of specific subgroups, we cannot limit ourselves to collecting and analyzing only health data, but we must take into consideration what comes from the world of social networking and smartphone apps, despite all the known limitations and criticalities.

**The evolution of the role of healthcare professional**

The exponential increase in knowledge and technologies, the greater complexity of the tasks, the increasingly diversified needs of the patients are flooding the healthcare professional with increasingly greater tasks and aspects to be addressed and managed. It is necessary to reflect on what the true essence of medical care is and to reassess the meaning of this profession. The ability to read change, juggle the new technology tools, exploit the potential of new IT and business analytics techniques will allow healthcare professionals to have greater support in their choices, to reduce the time spent collecting data or using the machines, to concentrate more on the decision-making process, rendering each intervention more effective and efficient.

**Data from Diabetology Literature**

Diabetology is facing different challenges: the ever decreasing number of diabetologists, the increasing number of patients, the reduction of visitation time, the increasing complexity of the pathology both from the clinical and welfare point of view, the difficulty in achieving the objectives, the growing burden of pathology management for healthcare providers and patients, accessibility to care and sustainability. New digital technologies and the use of *AI* are certainly a great opportunity. The current panorama of international scientific literature with respect to the use of Big Data and *AI* in Diabetology offers various hints and in-depth analyses, applicable to different fields.

**The epidemiological area**

With the identification of diabetes cases within large heterogeneous databases and the identification of new risk factors for diabetes. An interesting multi-database retrospective study on the identification of cases of T2DM [6] strategically used the EMIF Project database, a European project for the efficient re-use of health data for epidemiological research [7]. This database collects health information on about 52 million European citizens, using heterogeneous sources and acting as a support for the execution of high quality multinational observational studies, based on populations with sample sizes and otherwise inconceivable follow-ups. Using a complex algorithm strategy in eight different health data sources, subjects with T2DM were identified, and the strengths and limitations of each data source were revealed during the creation of a model that ensures the interoperability of systems of heterogeneous electronic medical records (EHR). These efforts represent a methodological advancement for carrying out studies of multi-national and multi-data sources, providing sufficient information for the contextualization and correct interpretation of the results and generating transparent and reusable documentation.
Additionally, a study conducted by researchers from the University of New York and Philadelphia described a new “data driven” approach of Population Health Management, based on the use of Machine Learning techniques to develop predictive models and identify risk factors for the onset of T2DM [8]. They based this model on administrative data concerning health services, pharmaceutical records, insurance databases, citizens’ access to the different care facilities and on laboratory results, from 4.1 million individuals for a period of 4 years, and a total of 42,000 variables. This model has been able to identify new risk factors for the appearance of diabetes with a predictive probability of at least 50% higher than the models based on known traditional risk factors, avoiding the costs of performing a screening.

The identification, through new predictive models based on Machine Learning, of the part of the population most at risk of developing diabetes will be able to generate clinical hypotheses for the identification of new risk factors and to implement more targeted cost-effective interventions.

**Phenotyping and risk stratification**

A study of 65,000 newly diagnosed T2DM patients [9] estimated that 10% of subjects absorb 68% of resources. The task of predictive analytical methods based on Big Data is precisely that of identifying the 10% of the subjects most at risk, on whom to intensify the treatment in order to improve the health outcomes at lower costs.

An interesting Japanese study [10] has developed a practical framework for phenotyping T2DM, using both specialized knowledge and an approach based on Machine Learning (in particular, the SVM - Support Vectors Machine) to develop, with data extracted from electronic medical records, two phenotyping algorithms: one with high sensitivity for screening, the other to identify subjects for research. Both algorithms showed superior performance compared to the basic algorithms, thus suggesting the possibility of using the proposed framework to conduct appropriate research depending on the objective.

**The diagnostic area.**

The use of predictive models based on AI has demonstrated the feasibility of identifying individuals with the highest probability of having undiagnosed diabetes, through clinical data that can be easily obtained from different databases [11], exploiting the potential of Machine Learning algorithms, including neural networks, as tools for diagnosing diabetes [12]. Machine Learning and pattern recognition are tools of enormous interest as they are promising in improving the sensitivity and/or specificity of disease detection and diagnosis, and appear to be able to reduce the potential for human error in decision-making. An exemplar application is the creation of surveillance algorithms able to detect diabetes and in particular to distinguish T1DM from T2DM by using structured electronic medical records [13]. The extraction of data from the electronic database of a 4-year long, multi-sectorial and multi-specialist outpatient practice allowed the inclusion of approximately 700,000 patients. Possible cases of diabetes have been reported using laboratory test results, diagnostic codes and prescriptions. Taking advantage of the entire range of data acquired from the records, more cases of diabetes have been captured compared to the analysis of administrative databases only, thus increasing the sensitivity of the method. The application of these algorithms to electronic folders has the potential to provide timely and clinically detailed information on large numbers, at low-cost, and nearly in real-time. Electronic records will probably become increasingly important sources for the surveillance of public health and for the definition of more targeted health policies [14,15].

**The field of automatic reporting**

Concerning the chronic complications of the disease, in particular Diabetic Retinopathy. The
automated classification of Diabetic Retinopathy has potential benefits such as increasing the efficiency, reproducibility and coverage of screening programs, reducing obstacles to accessing and improving results, providing early diagnosis and treatment. To maximize the clinical utility of the automated classification, an algorithm designed to detect specific lesions or to predict the presence of any level of Diabetic Retinopathy was recently developed [16]. In this study, Deep Learning, together with visual and pattern recognition techniques, has allowed to identify the desired features with the highest predictive value directly from the images, on a large set of labeled example data. These results show that deep neural networks can be "trained", using large data sets and without having to specify lesion-based features, to identify Diabetic Retinopathy or Diabetic Macular Edema in fundus images with high sensitivity and high specificity.

The economic field

Consisting in cost-effectiveness studies of health interventions. An Italian group performed a retrospective analysis through the cross-examination of large clinical and administrative databases with the aim of quantifying the relationship between health care costs attributable to diabetes and the level of control [17]. The results indicated that glycemic control (expressed by HbA1c levels) is a useful surrogate not only to estimate the odds of developing diabetes-related complications, but also for the costs associated with healthcare. The integration of administrative and clinical databases seems to be suitable to show that an appropriate management of diabetes can allow a better allocation of resources.

The expected advantages in diabetology

The ongoing medical revolution is strongly linked to the spread of digital health, the new software for AI, the use of big data to make more appropriate data-driven decisions, the even greater focus on predictive, preventive, personalized and participatory medicine. Each of these elements has important repercussions in the management of a complex and widespread chronic disease such as diabetes. Table 1 below summarizes the advantages and weaknesses attributed for each area of application, especially taking into account the challenges that Diabetology faces.

| ADVANTAGES | ISSUES |
| --- | --- |
| Digital Data management (glucometers and CGM connected to the cloud, data integration platforms) |  |
| • Support for doctors' decisions | • Difficult integration with computerized clinical records |
| • Reduced analysis time | • Different software for different devices |
| • Graphs and images easy to understand and interpret | • Time spent learning the software and gaining experience |
| • Correct management, supported by data, even remotely | • Risk of data "flooding" the professional and the patient |
| • Sharing with caregivers or family members | • Lack of significant evidences on the improvement of the outcomes |
| • Simultaneous analysis of data from different devices | • Limited number of patients currently accessing this technology |
| • Integration of glycemic values with alternative data for a better understanding (carbohydrate intake, physical activity) | • Lack of recognition for time spent and medical services |
| • Possibility to intervene in the intervals between visits | • Requirement of organizational changes |
| • Overcoming of geographical barriers |  |
| • Motivational tool |  |
| Mobile app (medical device with CE marking) |  |
| • Therapeutic instrument (eg bolus calculator) | • New skills and time for patient training |
| • Easier visualization of data and management of corrective actions | • Reliability of the instruments |
| • Overview of trends over time |  |
| • Greater patient involvement |  |
| • Facilitation for the patient |  |
| • Motivational and educational support tool |  |
| Telemedicine |  |
| • Overcoming of geographical barriers | • Non-recognition of medical services |

Table 1. Advantages and criticality of the use of new technologies in diabetology.
The Importance of the Chronic Care Model

A very interesting field of application concerns the Chronic Care Model (CCM), a medical assistance model for patients suffering from chronic diseases. The CCM is a model for sustainable chronicity [18], in which the scope is to achieve a "personalized and effective care", with an active participation of the person, integrating the different professionals involved in assistance, to improve concretely the life of the person with a chronic pathology. The CCM in the declination of its various dimensions (healthcare organization, delivery system design, decision support, clinical information systems, self-management support and community resources) brings on a huge amount of the so-called “Real World data”. Such Real World Data are data collected in the normal clinical practice (not from controlled clinical trials) which allow describing the patient's care pathways through the integration of different sources, consistent with what usually happens in reality. Each step of the care and assistance pathway, from diagnosis to treatment and follow-up, generates a large amount of data and images (Big Data) that often reside within health facilities in separate and independent databases. To obtain an integrated view of diagnostic-therapeutic pathways for individual patients and trace their complexity, it is necessary that these flows are integrated.

An interesting Australian study on CCM [19] shows how the use of electronic medical records and informatic integrated assistance can improve the management of chronic diseases. There are benefits for healthcare professionals and service users through an accurate and timely exchange of information, better work efficiency avoiding the repetition of data and information collection, as well as a better decision-making process.

The Perspective Of AMD

The Mission of the Italian association of Medical Diabetologists (AMD) is to promote the professional development of its members and to ensure the continuous improvement of the quality of assistance to all people with diabetes. In ancient times, AMD had already grasped the importance of standardizing the recording of daily clinical work in an electronic folder, followed by the collection and processing of real life information on care, examinations, complications and therapy, for a qualitative interpretation of real assistance on the national territory. This intuition of the use of variables with an intrinsic clinical and professional integrated meaning, which imply the doctor's reasoning in predictive and prescriptive choices - has in fact realized a first model of decisional learning thanks to the use of statistical algorithms.
The Experience of the Annals

The AMD Annals represent a periodic publication that has allowed, annually since 2006, to assess the care profiles of people with T1DM and T2DM, admitted to Italian diabetes-care centers. More in detail, a wide network of diabetology services equipped with a computerized clinical record system, used for the normal management of patients in charge, has a software provided by AMD that allows the extraction of a standardized set of clinical information (Data File AMD). The database obtained is used to calculate the quality of care indicators, both centrally and locally. In this way, a benchmarking activity is promoted based on the comparison of one's own performance with that recorded at national level (best performers approach). This activity reflects the professional performance of diabetologists, it is perfectly in line with the recommendations of the National Diabetes Plan, is very cost-effective, and has produced a systematic improvement over the years of all the indicators considered [20].

In addition, the AMD Annals database represents a valuable source of observational research data. In fact, it has allowed over time to investigate many key aspects, such as the care of the elderly patient, gender medicine, cardiovascular, renal and hepatic aspects, the appropriateness of the use of drugs and glycemic self-control, providing a realistic picture of the work of the Diabetology Services. Notably, the analysis of the Annals database has allowed the identification of critical areas and, therefore, the timely activation of processes of improvement, in a logic of Continuous Quality Cycle. It is easy, at this point, to visualize that the Annals project “photographs” the history of the evolution of clinical diabetology, and has allowed, for its own conception and structuring, each individual center to self-evaluate and improve in real time. Precisely this careful measurement of behaviors and results is capable of triggering a real dynamic evolution of an entire class of professionals, representing an irreplaceable benefit.

With an innovative and far-sighted thinking, AMD has created this database, which today represents an unparalleled heritage in a cultural civilization that is increasingly aware of the importance of information and the value of a large and accurate collection of data. In this new era, the world is being organized to take advantage of ever-larger databases to rely on technology and use them as assets to be interpreted to facilitate and accelerate important decisions in every field. It is clear that it is increasingly important to have a quality data collection that is increasingly “clean”, and it is fundamental to highlight the need for all professionals to be trained in the culture of the data and its correct recording. In fact, given that decisions can be made on the results of data processing, it is essential to raise the problem of the truthfulness of the data on which the analyses are based, to have a measure of the reliability of the results and be aware of them. Also in this area, AMD has invested time and resources in activating a process, born more than ten years ago, of culture on the quality of the data, which involved many partners who participated in the Annals project. This is why AMD feels ready to investigate the topic, with an advantageous background, and wants to project itself, in a competent and proactive way, in the world of Big Data and AIs, that represent the new cultural challenge also in the scientific field.

DIA&INT, the Value of the Diabetologist: Structured Skills, Prioritized Activities

The DIA&INT (Diabetes Intelligence) is an AMD project which rises in the setting of exploring new methods of analysis on unconventional data, and it can be considered a first experiment in analyzing unstructured data using the Business Intelligence method (a precursor of Machine Learning systems).

In a comprehensive vision of clinical activity, AMD focused on the value of the skills that identify the diabetologist and make it a decisive tool in the care process. To achieve this objective, AMD implemented a Business Intelligence project called DIA&INT [18]. This project is mainly aimed at encouraging the implementation of the Chronic Care Model in an “evidence based” way, by
highlighting the direct link between the activities carried out by the diabetologist and their expected outcomes, in order to highlight those activities which are essential to obtain the best outcomes, in a modern diabetology. Furthermore, this method may contribute to extrapolate the actions that could optimize the scarce resources, and may represent a valid support for the institutional choices in the current health system revision.

Admittedly, DIA&INT has been designed to respond to the need of establishing the qualitative dimension of the diabetologist's performance, and the specific contribution of the many factors participating to the clinical decision in the real world. To do this, AMD has chosen to use accredited tools:

1) organizational analysis: to measure and enhance the role of diabetes care with specific tools (SROI, Social Return On Investment);
2) data management with advanced tools (Business Intelligence).

The resulting “program” has been structured to standardize the information and define activities and competences, as implicitly described in official guidelines, measurable and comparable, with processing methods different from those of classical epidemiology, but necessary to perform predictive and prescriptive assessments. The methodology used [18] exploited the collective intelligence of the diabetologists who participated in a survey, expressing their opinion on a complex node such as the definition of the integrated benefit of certain activities in a personalized and weighted way. The intention was to “display” how diabetologists think when carrying out clinical judgment.

Through this approach two main concepts were outlined: the requirements (priority, specificity, frequency and multidimensional analysis), and the decision-making elements for “if…then” choices based on a priori knowledge (medical, regulatory, ethical, psychosocial, etc.), and the ability to propose a dynamic personalized treatment project to the person with diabetes that does not yet involve healing, and that requires an active personal effort.

By doing so, DIA&INT has selected the necessary skills and activities in medical practice, which have an impact on the evolution of the quality of the “health product” and the outcomes. DIA&INT produced the Core Competence curriculum of the diabetologist [21] and measured the impact of the activities on the experiences [18].

**Discussion**

The changes offered by technological innovation have generated an unprecedented level of data collection and processing, destined to undergo further expansion with the new applications of robotics and augmented reality, crossing a new frontier, and entering the era of Big Data and cognitive systems. A new category of technologies was born, which uses natural language processing and Machine Learning and is able to amplify and accelerate the digital transformation process, to allow people and machines to interact in a more natural way, extending and enhancing cognitive skills and abilities. The possibility of extracting information that has a meaning and is functional, in fact, requires the development of sophisticated technologies and interdisciplinary skills to operate closely together. In particular, in medicine, healthcare systems require consistent, appropriate and sustainable choices.

Today, the complexity of medicine certainly goes beyond the capacity of the human mind, the patients themselves are increasingly complex and we know that the long-term effectiveness of a treatment depends on variables that are no longer just “numerical”, but also on other information that is difficult to structure[22]. In this framework, advances in computing power play a central role in the acquisition of knowledge. It is essential to collect and use the key information in a coherent way, out of its abundance, by the means of effective and reliable analysis tools, that are represented by the new techniques of AI already available today. These techniques recognize and use Machine Learning systems that are able to “extricate themselves” and learn from these immense amounts of data, even
with intrinsic systems of recognition and error management. In essence, the AI is a machine capable of solving problems and reproducing activities typical of human intelligence [23].

AMD has already set the basis of culture and tools on the importance of collecting data, but now it is necessary to go further. The strong potential of the large amount of data collected in more than 10 years of observation in the AMD Annals Initiative may contain a significant “hidden” knowledge that, analyzed through different tools, may show new patterns that can help us to make better decisions in the prevention and management of diabetes and its complications. For example, this type of analysis may not only amplify our notions on the risk factors for diabetes complications, but it may also unravel new and unsuspected connections between them, indicating the probability with which they can affect the evolution of the disease on specific groups of subjects. Properly “trained” Machine Learning algorithms are able to efficiently evaluate millions of data while seeking probabilistic risk correlations, not limited to tracking the spread of an epidemic, but with the scope of identifying new personalized therapies. We could also imagine to go further, in the increasingly realistic hypothesis of linkage between the various healthcare databases (administrative data, assistance process, intermediate and final results, costs, etc.), which will allow us to make comprehensive assessments of the whole process of care in an even more individualized way in terms of clinical effectiveness, organizational effectiveness, sustainability, equity.

The field of Big data is already a reality and this systematic revision and position paper was made to offer the basis of knowledge for an constructive debate on the next steps that the AMD as an association and more in general Diabetology would follow in the very next future. It is time to choose to be proactive players in this complex system and take responsibility for these new processes aimed to improve the care of people with diabetes, which is and remains our “mission”.

Conclusions

In conclusion, the incessant progress of these changes brings into question many established paradigms, also in the scientific sphere, and this Position Statement represents a first document of careful analysis of this new world, which we must make our own in a logic of constructive comparison. We must reflect on the current scenarios, ask ourselves about the effects produced by these transformations to understand the consequences on our lives induced by the automated decisions and become able of integrating the traditional route into the new systems.

The diabetologist, more than any other healthcare professional, already has the right mentality, ready to take up this innovation: the culture of the data is in our DNA, the need to phenotype the patient and personalize the care and therapeutic approach have long been our priorities, the skills to manage a complex disease such as diabetes have been refined over time, ranging from technological to communicative, from educational and andragogical to managerial abilities. Thanks to all this progress made over the years and strongly desired by AMD, we are ready for a new challenge in the management of Diabetology that must see us as protagonists. The “digital diabetes” is coming, AI and big data are opening a window on new scenarios and today’s diabetologists must acquire new skills to be able to lead these changes, proactive in exploiting their potential and advantages, limiting their risks and guarding the essential elements for our profession of doctors and chronicity professionals.

Acknowledgements

N.M. conceptualized and drafted the manuscript. A.D.M. directed the “BIG DATA and AI“ Position Statement group of AMD, co-ordinated the writing of the manuscript and reviewed it. A.G., G.G., A.O. M.A.P., P.P., G.T.R., and R.Z. contributed by writing, editing and reviewing the manuscript. No funding was received for the preparation of this manuscript.
Conflicts of Interest

none declared

Abbreviations

AI: artificial Intelligence
AMD: Italian association of medical diabetologists
CCM: chronic care model
DIA&INT: diabetes intelligence
T1DM: type 1 diabetes mellitus
T2DM: type 2 diabetes mellitus
References

1. C. for D. and R. Health, Digital Health, FDA. (2019). http://www.fda.gov/medical-devices/digital-health (accessed October 21, 2019).

2. V. Iyengar, A. Wolf, A. Brown, K. Close, Challenges in Diabetes Care: Can Digital Health Help Address Them?, Clin Diabetes. 34 (2016) 133–141. PMID:27621530.

3. Ridi A. La gestione dei Big Data: strumenti di governance e appropriatezza. http://aemmedi.it/wp-content/uploads/2017/07/08_Europa_ridi.pdf., (n.d.).

4. P.W. Franks, E. Pearson, J.C. Florez, Gene-environment and gene-treatment interactions in type 2 diabetes: progress, pitfalls, and prospects, Diabetes Care. 36 (2013) 1413–1421. PMID:23613601.

5. S.K. Das, Integrating transcriptome and epigenome: putting together the pieces of the type 2 diabetes pathogenesis puzzle, Diabetes. 63 (2014) 2901–2903. PMID:25146470.

6. G. Roberto, I. Leal, N. Sattar, A.K. Loomis, P. Avillach, P. Egger, R. van Wijngaarden, D. Ansell, S. Reisberg, M.-L. Tammesoo, H. Alavere, A. Pasqua, L. Pedersen, J. Cunningham, L. Tramontan, M.A. Mayer, R. Herings, P. Avillach, P. Egger, R. van Wijngaarden, D. Ansell, S. Reisberg, M.-L. Tammesoo, H. Alavere, A. Pasqua, L. Pedersen, J. Cunningham, L. Tramontan, M.A. Mayer, R. Herings, P. Coloma, F. Lapi, M. Sturkenboom, J. van der Lei, M.J. Schuemie, P. Rijnbeek, R. Gini, Identifying Cases of Type 2 Diabetes in Heterogeneous Data Sources: Strategy from the EMIF Project, PLoS ONE. 11 (2016) e0160648. PMID:27580049.

7. EMIF, (n.d.). http://www.emif.eu/ (accessed October 21, 2019).

8. N. Razavian, S. Blecker, A.M. Schmidt, A. Smith-McLallen, S. Nigam, D. Sontag, Population-Level Prediction of Type 2 Diabetes From Claims Data and Analysis of Risk Factors, Big Data. 3 (2015) 277–287. PMID:27441408.

9. Comparative effectiveness for oral anti-diabetic treatments among newly diagnosed type 2 diabetics: data-driven predictive analytics in healthcare: Health Systems: Vol 2, No 2, (n.d.). https://www.tandfonline.com/doi/abs/10.1057/hs.2012.20 (accessed October 21, 2019).

10. R. Kagawa, Y. Kawazoe, Y. Ida, E. Shinohara, K. Tanaka, T. Imai, K. Ohe, Development of Type 2 Diabetes Mellitus Phenotyping Framework Using Expert Knowledge and Machine Learning Approach, J Diabetes Sci Technol. 11 (2017) 791–799. PMID:27932531.

11. Shankarakarayya, D. Odedra, S. Samanta, A.S. Vidyarthi, Computational Intelligence-Based Diagnosis Tool for the Detection of Prediabetes and Type 2 Diabetes in India, Rev Diabet Stud. 9 (2012) 55–62. PMID:22972445.

12. A.R. Olivera, V. Roesler, C. Iochpe, M.I. Schmidt, Á. Vigo, S.M. Barreto, B.B. Duncan, Comparison of machine-learning algorithms to build a predictive model for detecting undiagnosed diabetes - ELSA-Brasil: accuracy study, Sao Paulo Med J. 135 (2017) 234–246. PMID:28746659.

13. M. Klompas, E. Eggleson, J. McVetta, R. Lazarus, L. Li, R. Platt, Automated detection and classification of type 1 versus type 2 diabetes using electronic health record data, Diabetes Care. 36 (2013) 914–921. PMID:23193215.

14. A.N. Makam, O.K. Nguyen, B. Moore, Y. Ma, R. Amarasingham, Identifying patients with diabetes and the earliest date of diagnosis in real time: an electronic health record case-finding algorithm, BMC Med Inform Decis Mak. 13 (2013) 81. PMID:23915139.

15. L.M. Renard, V. Bocquet, G. Vidal-Trecan, M.-L. Lair, S. Couffignal, C. Blum-Boisgard, An algorithm to identify patients with treated type 2 diabetes using medico-administrative data, BMC Med Inform Decis Mak. 11 (2011) 23. PMID:21492480.

16. V. Gulshan, L. Peng, M. Coram, M.C. Stumpe, D. Wu, A. Narayanaswamy, S. Venugopalan, K. Widner, T. Madams, J. Cuadros, R. Kim, R. Raman, P.C. Nelson, J.L. Mega, D.R. Webster, Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs, JAMA. 316 (2016) 2402–2410.

https://preprints.jmir.org/preprint/16922 [unpublished, non-peer-reviewed preprint]
17. L. Degli Esposti, S. Saragoni, S. Buda, A. Sturani, E. Degli Esposti, Glycemic control and diabetes-related health care costs in type 2 diabetes; retrospective analysis based on clinical and administrative databases, Clinicoecon Outcomes Res. 5 (2013) 193–201. PMID:23696709.

18. N. Musacchio, R. C, idom, A. Cimino, A.D. Micheli, A. Giancaterini, L. Monge, A. Ozzello, M.A. Pellegrini, M.C. Ponziani, M. Ragonese, S.D. Cosmo, G.T. Russo, C. Suraci, R. Zilich, Diabetologist’s core competence curriculum: A position statement of the amd (Italian association of medical diabetologists), Diabetes Management. 9 (2019) 87–95.

19. S.-T. Liaw, J. Taggart, H. Yu, S. de Lusignan, C. Kuziemsky, A. Hayen, Integrating electronic health record information to support integrated care: Practical application of ontologies to improve the accuracy of diabetes disease registers, Journal of Biomedical Informatics. 52 (2014) 364–372.

20. Annali AMD – AMD, (n.d.). https://aemmedi.it/annali-amd/ (accessed October 21, 2019).

21. Musacchio N, Zilich R, Candido R, Cimino A, De Micheli A, Giancaterini A, Monge L, Ozzello A, Parodi S, Pellegrini MA, Ponziani MC, Ragonese M, Russo G, Suraci C., Core Competence Curriculum del diabetologo: Position Statement di AMD JAMD 20(2):S15-S32, 2017., (n.d.).

22. Z. Obermeyer, T.H. Lee, Lost in Thought - The Limits of the Human Mind and the Future of Medicine, N. Engl. J. Med. 377 (2017) 1209–1211. PMID:28953443.

23. Tullio De Mauro 2000 - Grande Dizionario Italiano Dell’Uso. Volume 6 - LibroCo.it, (n.d.). https://www.libroco.it/dl/Tullio-De-Mauro/UTET/isbn/Grande-Dizionario-Italiano-Dell-Uso-Volume-6/cw919033192436513.html (accessed October 21, 2019).