MEDINFO 2021: One World, One Health – Global Partnership for Digital Innovation
P. Otero et al. (Eds.)

doi:10.3233/SHTI220070

Design and Implementation of a Unique Patient Identification Model in Information Systems in Burkina Faso

Seydou Golo Barro[a,d,c], Adrien Ugon[b,c], Nadège R. Nana[d], Pascal Staccini[e]

[a] Nazi Boni University, Bobo-Dioulasso,Burkina Faso
[b] ESIEE-Paris, Cité Descartes BP99, Noisy-le-Grand, 93162, France
[c] Sorbonne Université, CNRS, LIP6, F-75005 Paris, France`
[d] Center for Training and Research in Medical Technology, Bobo-Dioulasso;
[e]RETINES Lab - IRIS department - Medical Faculty of Nice University, France

Abstract

The implementation of a reliable identity process is the basis of any secure patient information sharing system. Indeed, each individual is unique and should be identified by a unique number (identifier).

It is with these issues in mind that we have designed and implemented a unique patient identification method adapted to the context of Burkina Faso. The recommended method is inspired by the French method based on the work of the Group for the Modernization of the Hospital Information System (GMSIH) [1]. The developed model allows to assign a "Unique Identifier" (PatientID) to each patient from his profile of identification features (name, date of birth, gender,...). The patient ID is a sequence of 20 characters plus a security "key" of 2 characters. A reliability test of the model has been performed to take into account identity anomalies (duplicate, collision).

Keywords:

Unique Identifier, Patient Identification System, Hospital Information Systems

Introduction

The development of digital technology, and particularly of "communicating information technology", means that even in non-developed countries, we can now envisage the dematerialization of information exchanges on the patient. Indeed, more and more information is circulating between health actors for the care of the patient within health establishments and within health networks. The geo-political and health context, added to the challenges of digital technology, is forcing countries to increasingly computerize healthcare activities to facilitate and make communication between healthcare professionals and with the patient more reliable. The ideal in effective patient care is to provide care over increasingly long periods of time, by increasingly specialized professionals working in collaboration with each other in increasingly diversified healthcare organizations. Patient follow-up in these conditions requires fluid access to information wherever it is located and without mistaking the person. However, the diversity of identification systems within the hospital leads to increased risks of errors with enormous and sometimes fatal consequences. In the US, a collective of 27 organizations led by American Health Information Management Association (AHIMA) published an open letter to representatives on June 11, 2019, reporting that identification errors cost a healthcare facility an average of US$17.4 million (€15.5M) per year due to lost revenue and denied claims [2]. This justifies the establishment of a reliable patient identity. The reliability of the identification system is an essential condition for secure information exchanges. Indeed, patient identification is one of the basic functions of information systems. It allows the elements of the patient's record to be indexed by different systems and in different places, to ensure security, and to facilitate and make reliable access to the elements of the record by authorized health professionals and by the patient. Identifying the patient in the information system consists of assigning a new identifier or retrieving an existing identifier, based on these external characteristics called identification "features". Identification features are characteristics defined in an information system as constituting the identity of a patient and supposed to represent a person in order to uniquely identify him/her in this system. There are three types of features that constitute the "feature profile": strict features, extended features, and complementary features. These three types of features can be distinguished according to whether they are more or less easy to obtain, more or less stable, more or less confidential and more or less discriminating. The unique patient identifier is a sequence of numeric or alphanumeric characters. It is the "key" to access information about the patient. Even in developed countries, the electronic exchange of health data still raises many debates concerning patient privacy, the choice of patient identifier and the associated technical devices [3]. Indeed, for any sharing of patient information, regulations concerning patient rights and privacy must be applied. Likewise, precautions must be taken regarding the security and confidentiality of patient data. Appropriate procedures must also be put in place for the collection of patient identities and the procedures for reconciling these identities, with traceability of all operations. An identity surveillance unit must be set up. This unit will ensure that data is collected correctly, that anomalies and duplicates are reported, and that corrections are taken into account.

Status of the unique identifier in developed countries

The issue of unique patient identifiers has been raised in most countries of the world and each country has its own unique identification process. Patient identification is an important part of care management and is the first step in a process that lasts throughout the care journey. It is one of the important links in the health information system and contributes to the evolution of the health care system [4]. A study report by the Group for the Modernization of the Hospital Information System (GMSIH), «Patient identification principles and process» : International experiences with health systems and identification policy », presents the experiences of selected countries with patient identification. The report classifies countries into three broad categories [5].

- Countries where the national unique identification of the patient is organized or planned: Northern European countries (Denmark, Netherlands, Belgium, Great Britain, Ireland), New Zealand, and Germany.
- Countries where patient identification is more of a "regional" issue: Southern European countries (Spain and Italy), federal countries (Canada and Australia).
- Recent undecided countries: the United States of America moving towards the federation of identifiers

For the first category, there is Denmark where the patient identifier corresponds to the national identifier; for Belgium, the anonymized social security number is used as the patient ID; for Great Britain, the health patient ID "NHS number" is formed by the patient/doctor pair registered in the local NHS registry; in Ireland, the unique health patient ID is supported by a card called PPS (Personal Public Service) ; In New Zealand, the patient identifier is made up of 7 characters consisting of identification features; in Germany, each hospital has its own patient identification system made up of identification features;

For the second category: In Spain and Italy, the patient ID number is assigned by region. In Spain, the patient ID is based on a magnetic card. In Canada, the patient number is not standardized from one province to another. Each province assigns a health number to its inhabitants.

For the third category (undecided countries), notably the USA, after many discussions, the country is moving towards a federation of state identifiers (unique patient identifier at the national level) [2].

In France, the patient ID is obtained by combining several elements: the NIR (Number of Registration in the Directory of Natural Persons), the number of the health insurance card and the identification features.

Status of the unique identifier in Africa

In Africa, accurate patient identification remains a challenge for many health facilities due to the poor quality of information systems and the lack of national systems for identifying individuals [6] [7]. In Burkina Faso, the health system does not yet allow for end-to-end patient tracking and secure information exchange. The unique identification system is not yet used, which raises the problem of continuity of care when the patient returns to the system for other consultations. Initiatives are underway to set up a unique identification system for Burkinabé citizens and a unique national identifier [8]. In March 2018, a workshop was held in Ouagadougou under the auspices of the Ministry of Development of the Digital Economy and Posts to discuss the implementation of a unique electronic personal identifier in Burkina Faso [9]. It should also be noted that the General Directorate for the Modernization of Civil Status had a project to establish a unique identifier for citizens at birth. However, none of these initiatives has yet been completed. It is in light of all these issues that we decided to design a unique patient identification model in an African context and particularly in Burkina Faso.

Methods

Our model is inspired by the method used in France. Indeed, most of the identification features recommended by the GMSIH work are available in our context. The creation of the patient ID is done by an automatic algorithmic calculation method which is a process recommended in the GMSIH work. The information recorded at the time of the patient's admission is retrieved in the relevant fields to establish the identifier according to the defined protocol.

Creation of the Burkinabe model of unique identifier (PatientID)

A quality unique identifier must be built with reliable and always available features. Taken as a whole, these traits must allow a bijective relation between the physical person and his/her identification elaborated from these features. It should be remembered that there are three types of features that constitute the identification "feature profile". Strict features, extended features, complementary features. These three types of features can be distinguished according to their more or less easy way to obtain, more or less stable, more or less confidential and more or less discriminating character. Some features are more stable than others. These are retained to constitute the "strict features": generally names, first names, date of birth and sex. Features that are more unstable throughout a person's life, such as marital status or address, form the "extended features". They are used as discriminating elements of an administrative nature, when the strict features are insufficient to remove a doubt about an identity. Finally, the "complementary features" are health information, which allow professional actors to discriminate people even more finely. In our case, we have retained the following features: "Strict features": family name, first name(s), date of birth, sex. "Extended features": locality of birth. "Complementary features": Address of residence.

Criteria for creating the patient ID

In order to achieve a valid identification, it is necessary to follow a number of rules:

We have therefore developed an identification protocol to create a unique identifier in Burkina Faso that takes into account the local context. It is divided into four steps:

1. Use of identification traits. It consists in recovering the strict and secondary features of the patient necessary to the creation of the unique identifier;
2. Character standardization. This step allows to format the identification features. It consists of :
   - Standardize the first and last name fields in an adapted format (capitalization, no accents, no spaces for compound names);
   - Perform a first hash (using the sha512 hash algorithm) of the first and last name;
   - Code the gender field as 1 for male and 2 for female;
   - Normalize the date of birth to this format: ddmmyyy (8 digits);
   - Retrieve the birthplace code for the "place of birth" field. If the persons were born abroad the birthplace code is 00099 ;
   - Retrieve the residence location code for the residence address field. If the persons also reside abroad, the residence code is 00099 ;
   - Perform a concatenation on the result of the previous operations according to the following format:
   - (sex_code.ddmmyyy.hash(LastnameFirstname).birth_localisationcode.residence_localisationcode.area_num)
   - Keep only the digits of this concatenation and convert the letters into their corresponding integer according to the ASCII encoding
3. Hash: perform a second hash of the result obtained with the "sha512" algorithm, the goal being to reinforce the security of the identification system. After

this concatenation, take the first 20 characters from index 50.

4. Calculation of the control key: The control key allows to ensure the validity of the identifier. It is determined from the characters constituting the identifier code. By carrying out a Euclidean division by 97 of the numbers resulting from the hash, one obtains a remainder which will be subtracted from 97. The result of this subtraction is the control key.

Example of creation of a Patient ID

Let's take the following example of OUEDRAOGO Rachelle Evelyne, a patient born on May 14, 1991 in Koubri and residing in Sector 24 of Bobo-Dioulasso. We proceed to the creation of the patient ID by carrying out the steps below:

- Normalization and concatenation of the features “name”, “first names”: OUEDRAOGO Rachelle Evelyne.

Hash on the “name” and “first name”. This gives the result below:

b3360c85031e795a0980ebbec2ca8099df558d643bf804b28c06ea9eb62d69e316081672660fee709b0ec-dbdd96be1dce3b54ca2c23da74f5cdc076d3b301cfb

- Normalization and concatenation of the features “gender”, “date of birth”, “place of birth”, “home address”.

Result below:

219910514b3360c85031e795a0980ebbec2ca8099df558d643bf804b28c06ea9eb62d69e316081672660fee709b0ec-dbdd96be1dce3b54ca2c23da74f5cdc076d3b301cfb11894711081.

- perform a Hash on the result of the concatenation to obtain the following result:

(bb2389d2c3eb3c09774d11827d5148517d0c4222304abe610bd7bbe7554a1250f484286cc266a9b2294f03f9916efeb26fea9ce4a95651f5ad5cfead38b174b6) – Convert this result into ASCII format:

98985051565710050995110198519948575555521004949565055100534952565349551004899525050505148529798101544948981005598981015553535297495053481025256525056549999505454975798505057521024851102575749541011021019850541021019757991015297575354534910253971005399102101971005156984955529854.

- Choose the 20 characters from the conversion from an index to obtain the Unique Identifier corresponding to : 99102545055495049575
- Perform a Euclidean division by 97 of the numbers resulting from the conversion and subtract the result from 97 to obtain the key of the identifier, this gives 20.

Thus the patient ID of OUEDRAOGO Rachelle Evelyne is: « 99102545055495049575 20 ».

Practical application of the model in a health care system

Each patient is identified as soon as he/she comes into contact with the health care facility. The doctor collects information about the patient (features) from official civil status documents (identity card, passport, birth certificate, etc.). These features are used as criteria to search the patient in the information system (IS). There can be two cases: Case 1, the patient is not found in the IS. In this case, the doctor creates a new identifier by recording the information on the patient's identity (identification features). These are sent back to the program which automatically generates a patient ID according to the established protocol. Case 2, the patient already has an identifier. In this case, the doctor finds the patient's ID and accesses his file to complete it with the new consultation or hospitalization information. In fact, to find a patient whose ID exists, a search is performed in the identity index using the patient's identification traits as search criteria. The search for the patient is first performed on the “strict features” and the patient is either found or not found. The search on strict traits can also lead to a list of homonymous patients. In this case the search is extended to the extended features. The “extended features” are used when the strict features are insufficient to discriminate individuals. The search for complementary features is done as a last resort to remove doubt about an identity. It should be noted that anomalies can affect identifiers during registration operations. It is important to know how to identify and analyze these anomalies in order to take measures to correct them [10].

Identity anomalies

An “identity anomaly” in the hospital information system (HIS) occurs when the uniqueness of the identifier is no longer maintained. It can be: a “duplicate”, when two (or more) identifiers concern the same person; a “collision”, when the same identifier is attributed to two different persons (at least) figure 1. In both cases, it is necessary to extend the search on the trait profile in order to correct these anomalies. In the case of a duplicate, the Identity Vigilance Committee can proceed to merge the two identities. In the case of a proven collision, the correction could consist of identifying the type of collision, creating two new identifiers, establishing each of the two records under their new identifier.

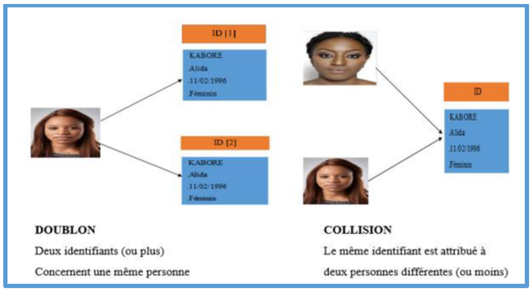

Figure 1

Results

Realization

To facilitate the use of the patient ID system we have developed an ID management application that can work with information systems. After the registration of patients in a HIS, it is possible to list their identity elements from a directory. From this directory, it is also possible to display, modify or delete a patient (figure 2).

Figure 2

Implementation of the patient ID system

For the implementation of the system, two types of libraries "***hashlib***" and "***unicode***" which allow respectively the hash function and the normalization of characters. The intermediate functions are "***Normalize***(string)" which uses the unicode library with the functions unicodedata.normalize and unicodedata.category. This function returns the string removed from capitals, accents, spaces, digits and special characters; the function ***Convers_ASCII***(string,L0=['0','1','2','3','4','5','6','7','8','9']) which takes as input a string and a list L0 of digits from 0 to 10. It decomposes the string into a simple list of elements using ***.strip*** ; it goes through this list, and converts the letters into integers using the function ***ord() ;*** it returns a concatenated list with the letters of the string converted into ASCII integers. The main functions are : 1) "***information_feature***()". It allows to enter the identification features of the patient. It returns a list of features concatenating the identification features entered : features = [last_name, first_name, date_of_birth, sex_code, birth_location_code, residence_location_code, area_number]. 2) « ***unique_id(features)*** » which takes as input the list of the previous features, then normalizes the name and surname fields, hash of the name and surname, first concatenation of the normalized fields, conversion of the letters of the hash into their ASCII integer, new hash, selection of the first 20 characters of the hash from index 50, re-conversion of the letters into their ASCII integer and selection of the first 20 digits, to give a first identifier, computation of the control key, concatenation of the identifier and the control key. It is also this function that returns the Unique Patient ID.

Integration of the unique patient identifier method in a HIS

The patient ID operates within a hospital information system. We test our model by integrating it into the electronic pregnancy monitoring system of the University Hospital of Bobo-Dioulasso (e-Lafia). E-Lafia is an information system for monitoring pregnancy. Patient information is collected in e-Lafia through a registration form. This information is stored in a MySQL database. The choice of the e-Lafia HIS is due to the fact that it uses health interoperability standards such as HL7v3, FHIR, DICOM, thus facilitating interaction with our unique identification system.

Procedure for integrating the patient ID system into the e-Lafia HIS

The integration of the system consists in writing a code allowing e-Lafia to recover the various features necessary for the automatic generation of the unique ID. Thus e-Lafia retrieves the ID calculation code in the form of a function. This calculation code generates the ID from the patient data collected in the patient registration form during the consultation with the doctor. To realize the procedure we have integrated our function in the HIS patient controller. This function first retrieves the information (features) needed to calculate the ID and then automatically generates the patient's Unique Identifier. Finally, it records the ID in the "patient" table of the e-Lafia database. The figure... represents the search interface of a patient already registered.

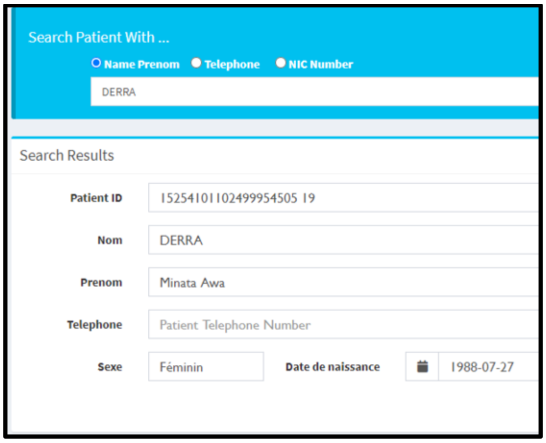

Figure 3

Quality testing of the patient ID system

It is recommended to regularly perform a quality assessment of the patient ID system in order to minimize the risk of anomalies (duplication and collision). To test the quality of our system we registered 212 patients in the e-Lafia HIS. The data was then exported from the database, specifically the patient table, in SQL format, which we converted to Excel format in order to perform statistical analysis of the data. To check for duplicates and collisions in our system, we used the duplicate detection method in Excel using the following formula reproduced in column F of the table in the figure 4.

=SI(NB.SI(E1:E1;E1)>1;"Doublon";"")

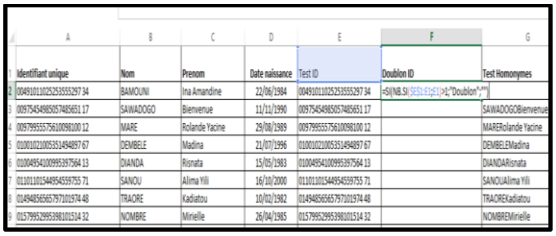

Figure 4

For the 212 records, the duplication and collision rate is 0%. We can therefore consider that our Unique Patient Identification system is relatively reliable. Indeed, the lower the indicator, the better the quality level of the identification system. [10].

Patient Identification Method Security Policy

The concept of security refers to all the hardware and software means implemented to reduce the risks of vulnerability of the patient identification system. The "sensitive" nature of the patient information exchanged requires the implementation of a strict security mode, which must comply with the following conditions [ ]. Thus, the following measures have been taken to ensure the security of our Patient ID system:

- Authentication, which ensures that access to resources (identification data) in the HIS is only authorized to entities (persons or machines) that are entitled to it. In short, it is a step that verifies the identity of the actor accessing the identification services.
- Data confidentiality: this is the protection of stored data against interception and reading by unauthorized persons. In our system, patient information is digitally coded. Thus, only authorized actors can see the information associated with each patient.
- Data integrity refers to the reliability and credibility of data throughout its life cycle. Thus, with the hash function in its 512-bit version, data integrity will be ensured.

Discussion

The objective of our work was to design and set up a unique patient identifier system in an African context mainly in Burkina Faso. We have produced a model which is inspired by the work of the Groupement pour la Modernization du System d'Information (GMSIH) in France [1]. This model was developed as a stand-alone application to allow it to work with any information system. To test the functioning of the model, we integrated it into an information system for monitoring pregnancy within the Souro Sanou University Hospital in Bobo-Dioulasso. The results of the evaluation show a duplicate and collision rate of 0%. The patient ID system will need to be improved to take into account as many possible situations as possible. Also, for any sharing of patient information, regulations concerning patient rights and respect for privacy must be applied. Likewise, precautions must be taken regarding the security and confidentiality of patient data. Appropriate procedures must also be put in place for collecting patient identities and procedures for reconciling these identities, with traceability of all operations. An identity monitoring unit must be installed. This unit will ensure the conditions for correct data collection, the reporting of anomalies and duplicates as well as the inclusion of corrections.

Conclusions

This pilot experience of setting up a unique patient identification system in Burkina Faso is an innovation and a foundation for computerization and patient data exchange projects. This work allowed us to realize once again the complexity of a health system and the challenges it represents. The work allowed us to examine the problem of identity vigilance in all its complexity. Based on the work of the GMSIH, we reflected on different scenarios of interactions between a patient and the healthcare system in order to select the most relevant identification criteria (features) possible. Although many things remain to be done before the system is operational in a fluid way, we can be satisfied to have contributed to the advancement of science for such a problem which remains topical in developed countries as well as in Africa, especially on the issues of confidentiality and privacy protection.

Acknowledgements

I thank all the collaborators who have contributed to the success of this work

References

[1] GMSIH, « Identification du patient : Guide pour l'élaboration des politiques d'identification et de rapprochement ». oct. 03, 2000.

[2] « USA : vers un identifiant patient unique au niveau national », *Innovation Mutuelle*, juin 18, 2019. https://www.innovation-mutuelle.fr/actualite/usa-vers-un-identifiant-patient-unique-au-niveau-national/ (consulté le mai 15, 2021).

[3] V. Dumont, « Controverses autour de l'échange électronique de données de santé. La question de l'identifiant du patient », *Termin. Technol. Inf. Cult. Société*, n° 108-109, Art. n° 108-109, juin 2011, doi: 10.4000/terminal.1347.

[4] « Accueil - Facturation - Recouvrement - Kit "Identification du Patient" ». http://ressources.anap.fr/afr/publication/607 (consulté le sept. 25, 2020).

[5] « Quantin CORTE - vs 13-07-07.pdf ». Consulté le: sept. 22, 2020. [En ligne]. Disponible sur: http://lertim.timone.univ-mrs.fr/Ecoles/infoSante/2007/supports_ppt/mardi%2017%20juillet/Quantin%20CORTE%20-%20vs%2013-07-07.pdf

[6] G. Bediang, E. E. Amougou, et M.-J. Essi, « Identification des Personnes Physiques et des Patients: Une Étude de Cas au Cameroun », 2019.

[7] F. Verbeke, G. Karara, S. V. Bastelaere, et M. Nyssen, « Patient Identification and Hospital Information Management Systems in Sub-Saharan Africa: A Prospective Study in Rwanda and Burundi », *Rwanda Med. J.*, vol. 69, n° 4, Art. n° 4, 2012.

[8] « Mise en œuvre de l'Identifiant unique de la personne au Burkina Faso : le comité de suivi et la Banque mondiale se concertent », *Ministère de l'Économie numérique, des Postes et de la Transformation Digitale*. https://www.mdenp.gov.bf/details?tx_news_pi1%5Baction%5D=detail&tx_news_pi1%5Bcontroller%5D=News&tx_news_pi1%5Bnews%5D=79&cHash=2393dbd3ee1b045cc1173595183eb7d9 (consulté le mai 15, 2021).

[9] « Burkina : Vers la mise en place d'une Identité unique de la personne - leFaso.net ». https://lefaso.net/spip.php?article82443 (consulté le sept. 30, 2020).

[10] GMSIH, « Identification du patient: Services d'identification du patient ,Tome 1 ». GMSIH, oct. 03, 2002.

[11] GMSIH, « Identification du patient: Principes de l'identification du patient Tome 1 », GMSIH, Paris, Groupe de travail, oct. 2002. Consulté le: sept. 25, 2020. [En ligne].

Address for correspondence

Dr Seydou Golo Barro
Institut Supérieur de Science de la Santé (INSSA)
Université Nazi Boni de Bobo-Dioulasso (UNB)
01 BP 1091 Bobo-Dioulasso 01, Phone: +226 70 46 88 98
Email : seydou.golo@gmail.com