DESIGN OF A MULTIVARIABLE PROGNOSTIC MODEL: A DEVELOPMENT AND VALIDATION STUDY TO PREVENT RE-STROKE IN PATIENTS WITH PREVIOUS CRYPTOGENIC STROKE

Universitat Autònoma de Barcelona, School of Medicine, Parc Taulí Teaching Unit

Author: Laura Gómez Dabó

Tutor: Dr. Antoni Martinez Rubio

Year 2019

INDEX

ABSTRACT

Introduction: Cryptogenic strokes represent approximately a 15% of all ischemic strokes and have a 3 to 6% recurrence risk. In most patients, low-risk sources of thromboembolism, such as a low-burden paroxysmal atrial fibrillation, are the presumed origin. A secondary prophylaxis treatment is not well-established yet.

Objective: To design a development and validation prediction model to stablish the probability of having a recurrent stroke (main objective) and mortality (secondary objective) in patients with cryptogenic stroke.

Methods: For the model development, a longitudinal 4 years follow-up retrospective cohort study would be done, in which 11 predictors would be combined statistically into a multivariable linear model with recurrent stroke as outcome. It would be followed by a temporal validation in a bigger sample.

Results: All predictors are expected to be risk factors except from anticoagulation treatment. These findings may help to find out more about the etiology of cryptogenic strokes and strategies to prevent them.

Keywords: prediction model · cryptogenic stroke · supraventricular arrhythmias.

RESUM

Introducció: Els ictus criptogènics representen aproximadament un 15% dels ictus isquèmics i tenen un risc de recurrència d'un 3 a un 6%. En molts pacients, fonts de baix risc de tromboembolisme, com la fibril·lació auricular paroxística, semblen ser l'origen. No hi ha cap tractament ben establert per la profilaxis secundària.

Objectiu: Dissenyar un model pronòstic de desenvolupament i validació per establir la probabilitat de patir un ictus recurrent (objectiu principal) i la mortalitat (objectiu secundari) en pacients amb ictus criptogènic.

Mètodes: Pel desenvolupament del model, disseny d'un estudi longitudinal retrospectiu de 4 anys de seguiment on es combinarien estadísticament 11 predictors en un model lineal multivariable obtenint ictus recurrent com a resultat; seguit d'una validació temporal en una mostra de major mida.

Resultats: s'espera que tots els predictors suposin un factor de risc a excepció de la anticoagulació. Aquestes troballes podrien ajudar a aclarir l'etiologia i possibles estratègies de prevenció d'ictus criptogènics.

Paraules clau: model pronòstic · ictus criptogènic · arítmies supraventriculars.

RESUMEN

Introducción: los ictus criptogénicos representan aproximadamente un 15% de los ictus isquémicos y tienen un riego de recurrencia de un 3 a un 6%. En muchos pacientes, fuentes de bajo riesgo de tromboembolismo, como la fibrilación auricular paroxística, parecen ser el origen. No hay ningún tratamiento bien establecido para la profilaxis secundaria.

Objetivo: Diseñar un modelo pronóstico de desarrollo y validación para establecer la probabilidad de sufrir un ictus recurrente (objetivo principal) y la mortalidad (objetivo secundario) en pacientes con ictus criptogénico.

Métodos: Para el modelo de desarrollo se realizaría un estudio longitudinal retrospectivo de 4 años de seguimiento donde se combinarían estadísticamente 11 predictores en un modelo lineal multivariable obteniendo ictus recurrente como resultado; seguido de una validación temporal en una muestra de mayor tamaño.

Resultados: Se espera que todos los predictores supongan un factor de riesgo a excepción de la anticoagulación. Estos hallazgos podrían ayudar a esclarecer la etiología y posibles estrategias de prevención de ictus criptogénicos.

Palabras clave: modelo pronóstico · ictus criptogénico · arritmias supraventriculares.

1. INTRODUCTION

Stroke is currently the second most common cause of death worldwide after ischemic heart disease (WHO, 2016) and the third largest contributor to disability-adjusted life-years (DALYs) from all cause in developed countries. The incidence and prevalence of stroke vary depending on factors such as geography, age, sex, ethnicity and socioeconomic status (1).

In Spain, stroke is the first cause of death in women and the third one in men (second cause in both gender) (INE, 2017) *(see Graphic 1)*. A recent study (IBERICTUS) observed an annual incidence for all cerebrovascular events of 187 new cases per 100.000 habitants. Incidence rates clearly increased with age in both genders, with a peak at or above 85 years of age. Of total of stroke patients, 81% were cerebral infarction, 16% were intracerebral hemorrhage, 3% were subarachnoid hemorrhage and 1% were unclassifiable stroke. In the subtype of ischemic stroke a 24 % was classified as undetermined cause (2).

Cryptogenic ischemic strokes (CIS) are defined as symptomatic cerebral infarcts for which no probable cause is identified after adequate diagnostic evaluation. Although the percentage of ischemic strokes that are classified as cryptogenic has declined over time from 40% in the 1970s as diagnostic testing has advanced, it stills represents a 10 to 15% even when extensive testing is applied in advanced centers (3). Cryptogenic stroke patients have a 3-6% risk of recurrent ischemic stroke, which is comparable to patients with other stroke causes (4).

From 80 to 90% of all CIS and an average frequency of 17% in global ischemic stroke, are recently classified as “Embolic stroke of undetermined source” (ESUS) which are defined as non-lacunar stroke on cerebral imaging and exclusion of large vessel atherosclerosis by CTA, MRA or ultrasound (5). In order to be classified as ESUS some criteria are required *(see table 1)*.

ESUS patients are younger, with a mean age of 65 years (42% women) with lower frequencies of conventional risk factors than non-ESUS patients. Their recurrence is rated in an average of 4.5% per year during a follow up of 2.7 years (5). Diverse low-risk sources are the presumed origin of thromboembolism causing infarcts in ESUS such as patent foramen ovale, aortic-arch atherosclerosis,

nonstenosing atherosclerotic plaques in cervical and intracranial arteries, mild left ventricular dysfunction, mitral annular calcification and low-burden paroxysmal atrial fibrillation (3).

There is a bidirectional relation between stroke and Atrial Fibrillation (AF). In one direction, AF is known to create a thrombogenic condition because of the stasis of the blood in the atria that can result in a stroke episode . In the other direction, it is well-established that strokes which affect different cerebral regions (such as the insular region, orbitofrontal and dorsal cingulate cortex, the hypothalamus, the amygdala, the periaqueductal gray or the ventrolateral medulla) may induce an autonomous imbalance causing a first episode of AF in stroke patients without known AF (4).

Several epidemiological studies have demonstrated that AF, which is the most prevalent arrhythmia worldwide (in Spain it affects from 1 to 2% of the population), multiplies, approximately, from 2 to 6 times the probability of having a stroke and 1,5 to 2,2 times the mortality. Strokes produced by AF not only cause high mortality but also disability and a higher tendency of recurrence compared with patients without AF. In fact, ESUS patients diagnosed with AF in the follow-up are older and more likely to have two or more infarcts in the same arterial territory in the initial magnetic resonance imaging. So that, AF is considered one of the prognosis determinants, being especially bad in women (6).

The increasing use of cardiac pacemarkers unveiled a long-surmised, but infrequently identified, group of patients with low-burden paroxysmal atrial fibrillation. In fact, a first episode of AF is detected in up to one third of cryptogenic stroke and in up to one fourth of ESUS patients with long-term monitoring. The frequency of detection is directly related to the duration of cardiac rhythm monitoring (5). The detection of AF after ischemic stroke is relevant for secondary stroke prevention (7), in which these patients are frequently switched from antiplatelet therapy to oral anticoagulation, although the relative benefits of it have not been well established (4).

There is little information about the association between supraventricular arrhythmias apart from AF and/or flutter and stroke. Specifically, the clinical relevance of short atrial runs, which are a frequent

finding after acute ischemic stroke, remains to be established and multiple studies have been performed to clarify which is its role (4).

In a recent doctoral thesis made by *Elisabet Pujol Iglesias M.D. Ph.D.*, named *Atrial Tachycardia and other predictor factors of recurrence of cerebrovascular events, atrial fibrillation and mortality in patients with cryptogenic stroke*, a study with 192 patients hospitalized for cryptogenic stroke was performed and results showed that patients without anticoagulation and with short atrial runs (24 hours Holter monitoring) had higher incidences of AF, recurrent stroke and cardiovascular mortality at 12 months of follow-up. After 4 years of follow-up, AF was still more frequent in patients with atrial runs, so that, patients with clinically suspected AF might benefit from anticoagulation (8).

Based on the above, it is evidenced that stroke is a major problem and that it is of great relevance for the doctor to easily predict the risk of recurrence and mortality of patients with stroke, based not only in personal characteristics (e.g. age and gender), but also on cardiovascular risk factors (e.g. chronical hypertension) and cardiological determinants (e.g. arial runs).

The main objective of this study is to develop and validate a clinical prediction model to stablish the probability of having a recurrent stroke in patients with a previous cryptogenic stroke in a 4 years period, based on their gender, age, presence or absence of chronical hypertension, dyslipidemia, diabetes mellitus, glomerular filtration rate, previous stroke, previous heart disease, left atrium size, atrial runs and anticoagulation treatment.

The secondary objective is to predict mortality of cardiovascular and non-cardiovascular causes based on the multiple explanatory variables at 4-years follow-up.

2. METHODS

2.1 Study design and source of data

This prediction model would be a longitudinal 4 years follow-up retrospective cohort study which would combine both categories, model development and model validation. It would be carried out

following the Tripod Statement (9). For its realization two datasets of cryptogenic stroke patients would be collected, one for each category *(see Figure 2)*.

For the model development, a sample would be obtained from the database created by Dr. *Elisabet Pujol Iglesias.* Specifically, 208 consecutive patients with a cryptogenic stroke diagnosis in their medical discharge signed by the Neurology Department of *Parc Taulí* Hospital dated between January 2010 and September 2013.

For the validation model, a sample formed by 400 patients would be recruited. This sample would be obtained by a sum of the previous 208 consecutive patients of the model development with 192 new patients. These new patients would be recruited from September 2013 onwards until reaching the total of 192 new consecutive patients with cryptogenic stroke diagnosis in their medical discharge signed by the Neurology Department of *Parc Taulí*. The same predictors, measurements and outcome definitions would be applied. This way a bigger sample and a temporal validation would be possible.

2.2 Participants

As it was aforementioned in *Source of data,* participants of the study would be recruited from one tertiary care, *Parc Taulí* Hospital of Sabadell.

Individuals would enter or not in the cohort study on the basis of specific criteria. The same inclusion and exclusion criteria would be used both for the development and validation of the model.

- Inclusion criteria: adult patients (age 18 years old and above) of *Parc Taulí* Neurology Department with diagnosis of cryptogenic stroke in their medical discharge in who electrocardiogram, echocardiogram and 24 hours Holter monitoring were practiced during their hospitalization.
- Exclusion criteria: patients hospitalized for hemorrhagic stroke or ischemic stroke with a defined cause in their medical discharge.

All patients would be asked to sign a written informed consent *(an example form is included in the addendum)*. The study would be evaluated by the Ethics Committee of the hospital.

2.3 Outcome

In order to respond to the main objective, the principal outcome of interest would be re-stroke, i.e. recurrent stroke. Recurrent stroke would be defined as one or more acute episode(s) of focal dysfunction of the brain, retina, or spinal cord lasting longer than 24 h, or of any duration if imaging (Computerized Tomography or Magnetic Resonance Imaging) or autopsy show focal infarction or hemorrhage relevant to the symptoms (10), in a person with a previous cryptogenic stroke.

In our study, as specified in the inclusion criteria, all patients would have suffered and been diagnosed of cryptogenic stroke. At the time of the episode, patients would enter the study and would be followed-up through their clinical hospital and primary care history for 4 years. In these years, the presence or absence of one or more stroke episodes after the date of inclusion, understood as recurrent stroke (s), would be documented *(see Figure 3)*. During the following-up, death from any cause and cardiovascular mortality would also be documented as a secondary outcome to reply to the secondary objective.

No actions to blind assessment for the outcomes would be applied because of their objective nature.

2.4 Predictors

Predictors would be obtained from the patient's medical records just at the time of entering the study. These predictors are objective, easy to collect and measure which would make this model applicable anywhere in the world.

The following data would be extracted for each patient: gender (male/female), age, cardiovascular risk factors (chronical hypertension, dyslipidemia, diabetes mellitus), renal function (Glomerular Filtration Rate), the presence of previous stroke, the presence of previous coronary heart disease, left

atrium size (measured by echocardiogram in long parasternal shaft), the presence or not of auricular runs (defined as presence of 4 or more consecutive ectopic auricular beats in the electrocardiogram) and anticoagulation treatment or not.

Regarding the treatment received, it is known that some patients with cryptogenic stroke are treated with anticoagulation because of a high suspect of AF although there is not firm evidence about it. Because it may influence the prognosis, it would be included as a predictor in the model development. If the predictive effect of the intervention is rather small compared with the other predictors, the treatment would be excluded from the modeling.

Due to the fact that predictors would have been collected from people unconnected with the study, that predictors do not require subjective judgment, and that the incremental value of each predictor would be quantified, no blind assessment of predictors would be considered.

2.5 Sample size

Sample size would be restricted to using an available dataset. As exposed previously in *source of data*, 208 patients and 400 patients would be included for the development and validation study, respectively.

2.6 Missing data

Instead of omitting all individuals with any missing value or using the missing indicator method, multiple imputation would be applied, i.e. multiple copies of the data set with the missing values replaced by imputed value drawn from the predicted distribution would be created by using the observed data.

2.7 Statistical analysis methods

This study intends to develop a model with only a few major predictors to increase clinical applicability.

For the model development, each predictor would define an explanatory variable. Gender (female/male), cardiovascular risk factors (chronical hypertension, dyslipidemia and diabetes mellitus), coronary heart disease, previous stroke, auricular runs and anticoagulation treatment would be handled as categorical variables (presence or absence). For continuous predictors (age, left atrium size and glomerular filtration rate) the presence of linear or nonlinear relationship would be explored with restricted cubic splines.

In order to respond to the main objective of the study, establishing the probability of re-stroke, all 11 predictors would be combined statistically into a multivariable linear model (quantitative outcome: probability of re-stroke after 4 years). The result would be the impact of each variable on the odds ratio of re-stroke. We would follow this formula:

$$\gamma_i = \beta_0 + \beta_1 X_1 + \beta_2 x_2 + \cdots \beta_m x_{\mathrm{m}}$$

Where γ_i indicates the probability of re-stroke, $ß_0$ is constituted by those individuals presenting the reference level of each and every variable $x_{1\ldots\mathrm{m}}$, $ß_i$ are the coefficients associated with the reference group and x_i are the explanatory variables.

For predictor selection during modeling, a backward elimination would be used to consider all correlations between predictors and the outcome. To avoid making an overfitted and optimistic model, knowing our sample size would be small, a higher value of predictor's significance would be considered.

We would assess internal validity with a bootstrapping procedure to quantify any optimism in the final prediction model and obtain a realistic estimate of the performance. Calibration (i.e. the

agreement between predictions from the model and observed outcomes) would be reported graphically. Discrimination would be obtained by measuring the concordance index (c-index).

To respond to the secondary objective of the study, predicting mortality 4 years after a recurrent-stroke episode, Cox proportional hazard regression model would be used. Hazard ratio would be calculated, and a Kaplan-Meier curve obtained.

Finally, for the model validation, for each individual in the new dataset, outcome predictions would be made using the published linear formula and compared with the observed outcomes.

3. EXPECTED RESULTS

This is a design of a study, so no results could be been obtained.

If the study was actually done, ß coefficients would be obtained with a standard error and a P value. P value would identify significant predictors, which would be included in the formula. Depending on whether the *ßi* coefficients result negative or positive, we would establish if the multiple explanatory variables are risk or protective factors, respectively.

Because of the information that it's actually known about the different proposed predictors, we expect all of them to be risk factors for re-stroke and mortality, except from the anticoagulation treatment, which we expect to be a protective factor.

After having all this information, a formula could be created to easily predict which is the probability of a patient who has suffered a cryptogenic stroke either to have a recurrent stroke and/or to die 4 years after his cryptogenic stroke episode.

4. IMPLICATIONS AND FUTURE INVESTIGATION

Cryptogenic stroke is still a major challenge, especially because as long as no specific cause is found, adequate or therapeutic measures cannot be applied, and with it the recurrence percentage is considerably high.

A risk score able to identify a high risk of recurrent stroke would be a major improvement for the management of these patients. Thus, knowing which predictors are strongly associated with recurrent stroke might help to find out more about its etiology, might precisely identify which patients might benefit from prolonged ECG monitoring after a cryptogenic stroke (which is now left at the physician's discretion) and a confident detection of high-risk patients may encourage to assume the risks of applying a more intensive treatment (such as anticoagulation).

Furthermore, with a reliable risk score, our medical health system, which currently spends from 3 to 4% of the total health care expenditures on stroke (1), might benefit as well reducing direct and indirect costs derivate of cryptogenic stroke. For example, unnecessary tests could be avoided in low risk patients, who could be visited in less expensive outpatient clinics.

Besides, avoiding re-strokes would result in reducing intangibles costs and in improving the quality of life of these patients by decreasing the disability associated with each stroke episode.

For future investigations, if the presence of auricular runs result to be a risk predictor and the anticoagulation treatment a protective predictor of having a recurrent stroke, a clinical trial could be designed to try to verify if the incidence of stroke reduces (or not) with its detection and treatment.

5. DISSEMINATION PLAN

If the study responds appropriately to the main and secondary objectives, a dissemination plan would be followed step by step *(see figure 4)*.

First of all, the study would be presented in the Cardiology Service of Parc Taulí Hospital, not only in order to explain our results, but also to debate the weaknesses and strengths of the study. With all the inputs of the different professionals we would write manuscript and prepare oral presentations for the next step.

Secondly, we would present those results and conclusions to local congresses such as, *Societat Catalana de Cardiologia (SCC)* Congress and *Sociedad Española de Cardiología* Congress (SEC),

and to other international congresses with the intention of diffusing it and contrasting similar or antagonist findings around the world. A manuscript would be also prepared for a specialized medical journal.

Thirdly, on one hand, in the interest of spreading the results within the Parc Taulí Hospital, we would program sessions for doctors of different departments where stroke has an important impact (such as Internal Medicine, Neurology and Geriatric Services). Besides, we would program conferences oriented to the hospital's patients who have suffered a cryptogenic stroke and to their families for giving them information about the new findings. On the other hand, for disseminating the results in the surrounding area, sessions for General Practitioners in their primary care centers would be carried out.

Finally, we would write a press release to raise awareness in local population about stroke and AF, to explain the new findings and how specialists in Parc Taulí Hospital are working to improve the actual and future situation of people affected by this disease.

6. BIBLIOGRAPHY

1. WHO R. Global burden of stroke. 2003;51.

2. Díaz-Guzmán J, Egido JA, Gabriel-Sánchez R, Barberá-Comes G, Fuentes-Gimeno B, Fernández-Pérez C. Stroke and transient ischemic attack incidence rate in Spain: The IBERICTUS study. Cerebrovasc Dis. 2012;34(4):272–81.

3. Saver JL. Cryptogenic Stroke supplement. N Engl J Med. 2016;374(21):2065–74.

4. K.G. H, S. T, R.B. S. Detection of Atrial Fibrillation in Cryptogenic Stroke. Curr Neurol Neurosci Rep . 2018;18(10):1–7.

5. Hart RG, Catanese L, Perera KS, Ntaios G, Connolly SJ. Embolic Stroke of Undetermined Source: A Systematic Review and Clinical Update. Stroke. 2017;48(4):867–72.

6. Martínez-rubio A, Iglesias EP, Thio MB, Mas EG. Epidemiología de la fibrilación auricular en España. 2015;13(C):3–8.

7. Diener HC, Bernstein R, Hart R. Secondary Stroke Prevention in Cryptogenic Stroke and Embolic Stroke of Undetermined Source (ESUS). Curr Neurol Neurosci Rep. 2017;17(9):17–21.

8. Pujol-Iglesias E. Taquicàrdia auricular i altres factors predictors de recurrència d'esdeveniments cerebrovasculars, fibril·lació auricular i mortalitat en pacients amb ictus criptogènic [dissertion]. Sabadell: Universitat Autònoma de Barcelona; 2018. 140 p.

9. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD Statement. Eur Urol. 2015;67(6):1142–51.

10. Mansfield A, Inness EL, Mcilroy WE. Stroke. Handb Clin Neurol. 2018;159(16):205–28.

7. ADDENDUM

7.1 Figures

Figure 1. Causes of death of both gender in Spain (INE, 2017).

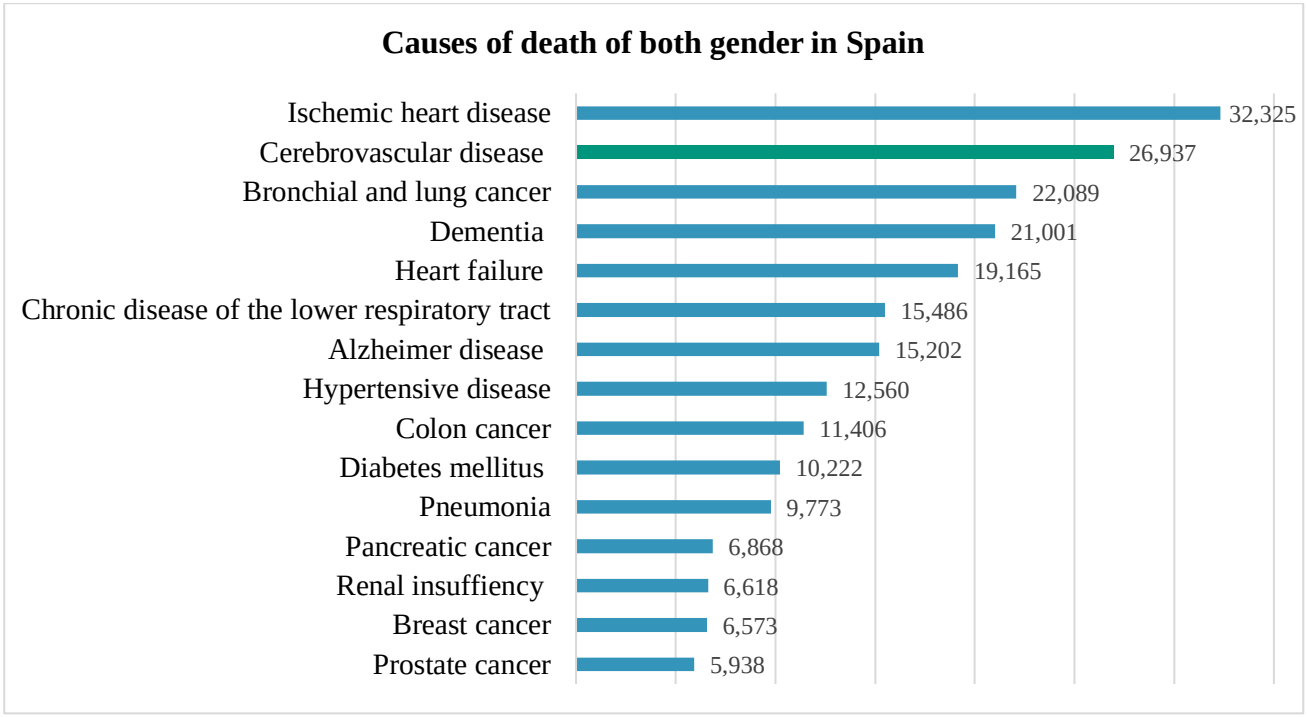


Figure 2. Schematic representation of the development and validation of the prediction model.

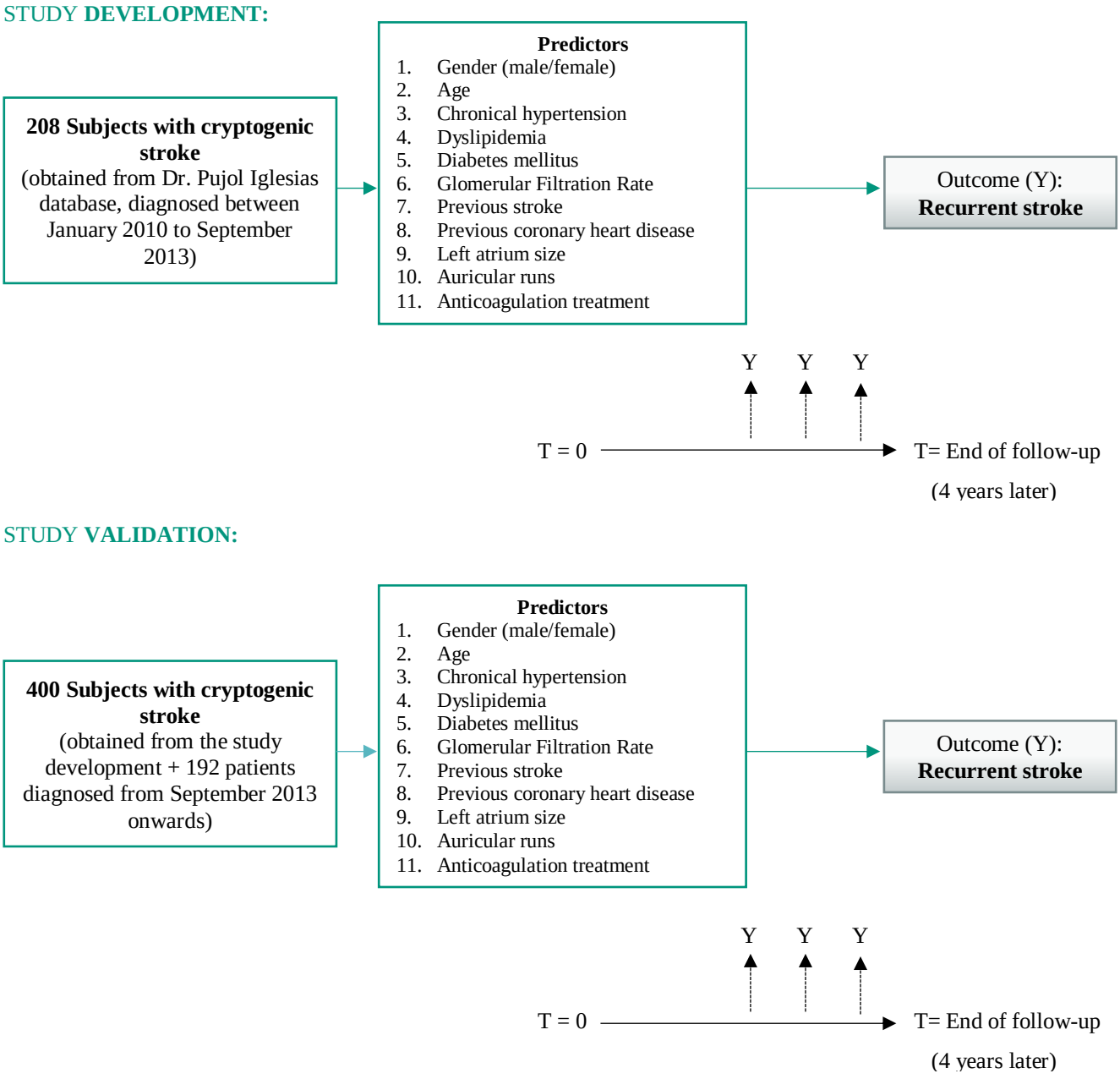


Figure 3. Example of follow-up.

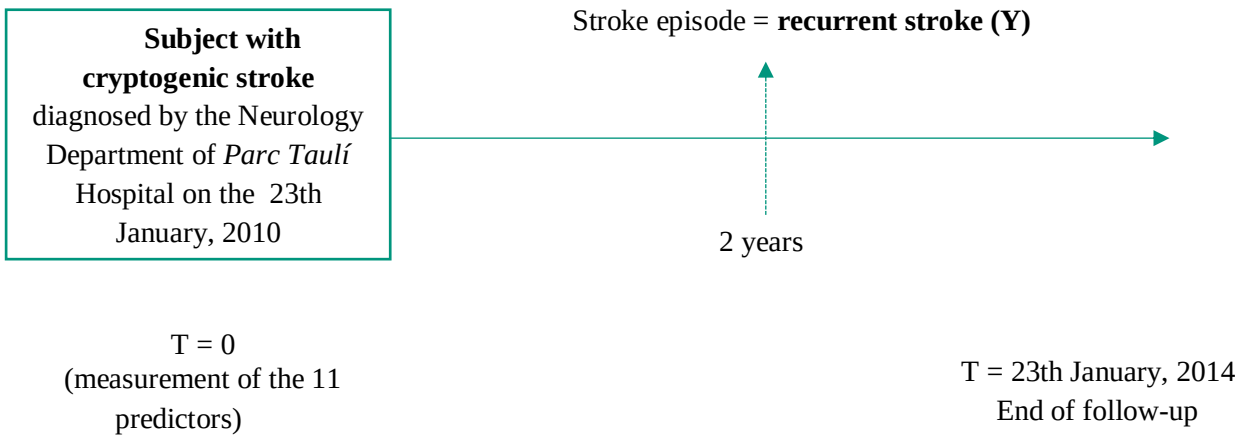
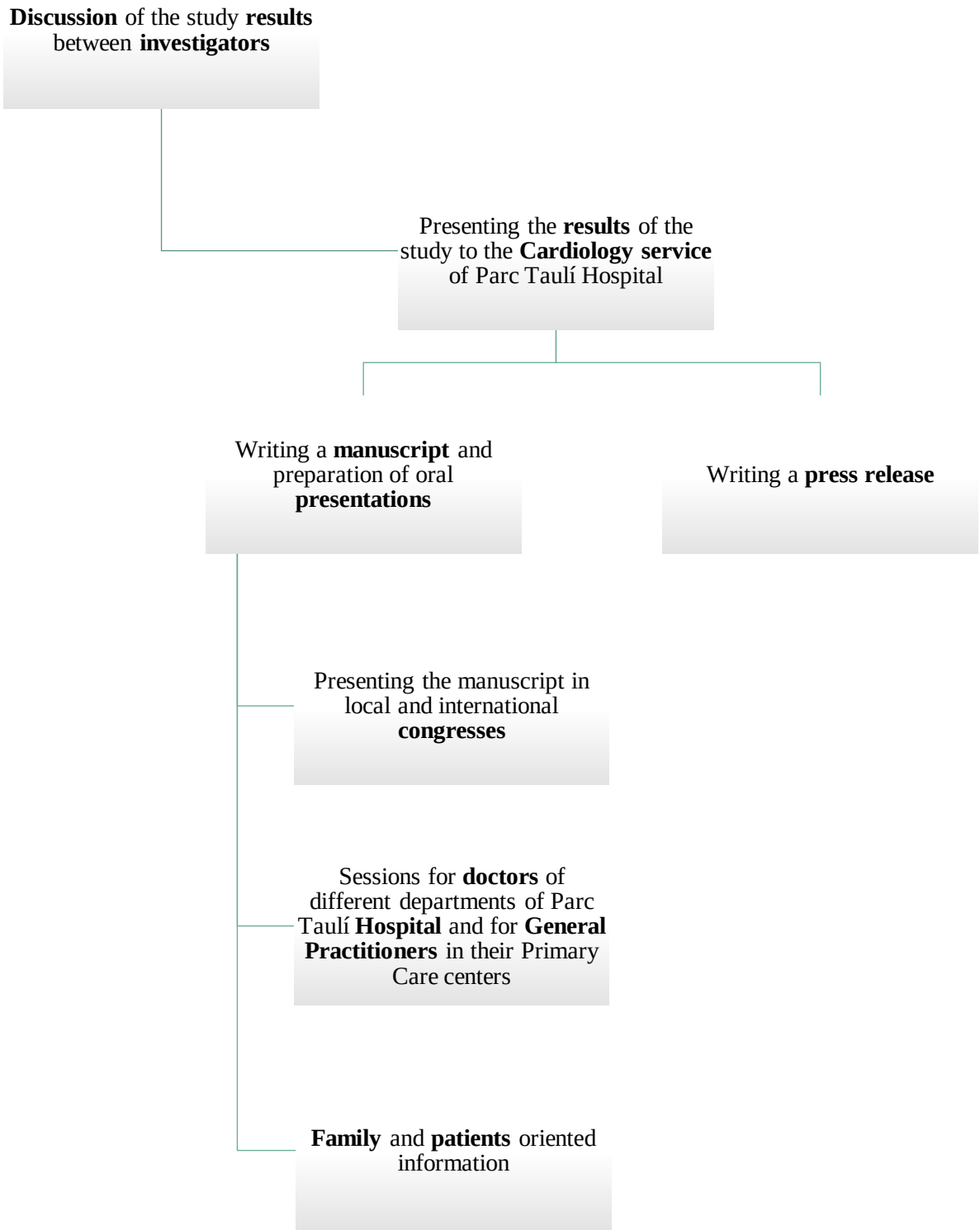


Figure 4. Dissemination plan.



7.2 Tables

Table 1. Criteria for Diagnosis of Embolic Stroke of Undetermined Source (ESUS)(5)*****

1.	Ischemic stroke detected by CT or MRI that is not lacunar†.
2.	Absence of extracranial or intracranial atherosclerosis causing ≥50% luminal stenosis in arteries supplying the area of ischemia.
3.	No major risk cardioembolic source of embolism✶.
4.	No other specific cause of stroke identified (e.g., arteritis, dissection, migraine/vasospasm, and drug abuse).

***Abbreviations**: CT:* computed tomography; *MRI:* magnetic resonance imaging.

*Requires minimum diagnostic evaluation that includes cardiac rhytm monitoring for >24 hours with automated rhythm detection.

† Lacunar defined as a subcortical infarct ≤ 1.5cm (≤2.0 cm on MRI diffusion images) in largest dimension, including on MRI diffusion-weighted images, and in the distribution of the small, penetrating cerebral arteries of the cerebral hemispheres and pons.

✶Permanent or paroxysmal atrial fibrillation sustained atrial flutter, intracardiac thrombus, prosthetic cardiac valve, atrial myxoma or other cardiac tumors, mitral stenosis, recent (<4 weeks) myocardial infarction, left ventricular ejection fraction<30%, valvular vegetations, or infective endocarditis.

7.3 Documents

INFORMED CONSENT FORM FOR PARTICIPATING IN THE STUDY: 'A MULTIVARIABLE PROGNOSTIC MODEL: A DEVELOPMENT AND VALIDATION STUDY TO PREVENT RE-STROKE IN PATIENTS WITH PREVIOUS CRYPTOGENIC STROKE'.

This Informed Consent Form is for patients (adults including men and women) who were diagnosed of cryptogenic stroke by the Parc Taulí Neurology Department from January 2010 onwards.

This study is carried out by Laura Gómez Dabó as a principal investigator, supported by the Parc Taulí Hospital. The name of the study is: 'A multivariable prognostic model: a development and validation study to prevent re-stroke in patients with previous cryptogenic stroke'.

This Informed Consent Form has two parts:

- **Information Sheet** (to share information about the research with you).
- **Certificate of Consent** (for signatures if you agree to take part).

You will be given a copy of the full Informed Consent Form.

PART I: Information Sheet

Introduction

I am Laura Gómez Dabó, a student of sixth year of Parc Taulí Hospital. We are doing a study to develop and validate a model in order to establish the probability of having a recurrent stroke after a cryptogenic stroke episode. I am going to give you information and invite you to be part of this study. You don't have to decide today whether or not you will participate in the study. Before you decide, you can talk to anyone you feel comfortable with about the study.

There may be some words that you do not understand. Please ask me to stop as we go through the information and I will take time to explain. If you have questions later, you can ask them to me, the study doctor or the staff.

Purpose of the research

Stroke is a major problem because is one of the most common cause of death and disability worldwide. Ischemic strokes are the majority of strokes (81% in Spain) and occur when blood don't arrive to different parts of the brain. If after the study a cause is not founded to be the origin, the stroke is called 'cryptogenic stroke', which are almost 1 out of 5 of all ischemic strokes. After having one stroke it stills remains a certain probability of having a second one. Although there are some determinants which are suspected to increment the risk of having a second episode, they are not well established. The problem is that if a cause is not determined, it is difficult to indicate a treatment to prevent another episode (what is formally known as prophylaxis treatment to prevent a recurrence).

The reason of this study is to establish, based on different determinants (specifically sex, age, cardiovascular risk factors, renal function, the presence of previous coronary heart disease, previous stroke, left atrium size, auricular runs and anticoagulation treatment), the probability of having a recurrent stroke and the survival in persons who have already suffered a cryptogenic stroke.

The reason why these determinants are chosen is because they would be applicable worldwide with easy questions and measurements. With the study we will know what weight each of the determinants has and a

formula will be created. The formula will permit an easy calculation of the probability of recurrence and survival allowing patients and their doctors to decide whether to consider the risks of taking or not a treatment.

Type of Research Intervention

This study will require you allowing us access to your medical records from the date of the cryptogenic stroke until 4 years after it.

Participant selection

We are inviting all adult patients from Hospital Parc Taulí who had a cryptogenic stroke diagnosed by the Neurology Department between January 2010 onwards to participate in the study.

Voluntary Participation

Your participation in this study is entirely voluntary. It is your choice whether to participate or not. Whether you choose to participate or not, all the services you receive at this Hospital will continue and nothing will change. You may change your mind later and stop participating even if you agreed earlier.

Procedures and Protocol

This study will require you allowing us access to your medical records from the date of the cryptogenic stroke. We will obtain information of the moment of the cryptogenic stroke episode, about:

- Your cardiovascular risk factors (if you had chronical hypertension, dyslipidemia or diabetes mellitus).
- Your renal function (with the Glomerular Filtration Rate).
- Previous coronary heart disease and/or stroke(s).
- Cardiological measurements during your hospitalization (left atrium size and the presence or not of auricular runs in the electrocardiograms).
- Whether or not anticoagulation treatment was prescribed.

We will also follow you through your medical records during the 4 years after the cryptogenic stroke episode to determine if there was a stroke recurrence or not and to stablish survival.

Risks

No risks are expected.

Benefits

As an individual, although the only benefit you will obtain would be applying the results of the study to your individual situation, your participation will help us to continue unraveling why one person with cryptogenic stroke suffers another one or if a second episode haven't occurred yet, which is the probability of happening. Having the probability in terms of numbers would help to consider applying a treatment, always considering the risk-benefit. There may not be any benefit to the society at this stage of the research, but future generations are likely to benefit.

Reimbursements

You will not be given any money or gifts to take part in this study.

Confidentiality

The information that we collect from this research project will be kept confidential. Information about you that will be collected during the research will be put away and no-one, but the researchers will be able to see it. Any information about you will have a number on it instead of your name. Only the researchers will know which is your number and it will be kept confidential. It will not be shared with or given to anyone.

Sharing the Results

The knowledge that we get from doing this research will be shared with you if you are interested. Once the study is finished, we can program a visit with us to explain it to you and apply, if you want it to, the results to your individual situation. At the same time, we will publish the results (confidential information will not be shared) in order that other interested people may learn from our research.

Right to Refuse or Withdraw

You do not have to take part in this research if you do not wish to do so. You may also stop participating in the research at any time you choose. It is your choice and all of your rights will still be respected.

Who to Contact with?

If you have any questions you may ask them now or later, even after the study has started. If you wish to ask questions later, you may contact Laura Gómez Dabó by email (laura.gomezd@e-campus.uab.cat).

This proposal has been reviewed and approved by the Ethics Committee of Parc Taulí Hospital which is a committee whose task it is to make sure that research participants are protected from harm.

PART II: Certificate of Consent

I have read the foregoing information, or it has been read to me. I have had the opportunity to ask questions about it and any questions that I have asked have been answered to my satisfaction. I consent voluntarily to participate as a participant in this research.

Print Name of Participant_____

Signature of Participant _____

Date _____

Day/month/year

If illiterate

A literate witness must sign (if possible, this person should be selected by the participant and should have no connection to the research team). Participants who are illiterate should include their thumb-print as well.

I have witnessed the accurate reading of the consent form to the potential participant, and the individual has had the opportunity to ask questions. I confirm that the individual has given consent freely.

Print name of witness_____ **AND** **Thumb print of participant**

Signature of witness _____

Date _____

Day/month/year

Statement by the researcher/person taking consent

I have accurately read out the information sheet to the potential participant, and to the best of my ability made sure that the participant understands that the following will be done:

1.

2.

3.

I confirm that the participant was given an opportunity to ask questions about the study, and all the questions asked by the participant have been answered correctly and to the best of my ability. I confirm that the individual has not been coerced into giving consent, and the consent has been given freely and voluntarily.

A copy of this ICF has been provided to the participant.

Print Name of Researcher/person taking the consent_____

Signature of Researcher /person taking the consent_____

Date _____

Day/month/year