A risk stratification using machine learning techniques to identify types of high-risk multiple chronic conditions patients, their needs and subsequent organization of integrated care services.

Una estratificación del riesgo utilizando técnicas de *machine learning* para identificar tipos de pacientes pluripatológicos de alto riesgo, sus necesidades y posterior organización de servicios de atención integrada.

Tesis Doctoral escrita por **Pablo Evaristo Bretos Azcona**

Dirigida por **Dr. Juan Manuel Cabasés Hita** & **Dr. Eduardo Sánchez Iriso**

Universidad Pública de Navarra (UPNA)

https://doi.org/10.48035/Tesis/2454/40067

Agradecimientos

A mis directores de tesis, Eduardo y Juan, por su guía y consejo durante estos años.

A Cristina Ibarrola y el Servicio de Efectividad y Seguridad Asistencial del Servicio Navarro de Salud-Osasunbidea por la oportunidad de colaborar en el programa de atención integrada para pacientes crónicos y pluripatológicos.

Al Gobierno de Navarra por su apoyo a través de una beca de investigación doctoral.

INDEX

EXECUTIVE SUMMARY

Background

Patients with multiple chronic conditions (MCC) are characterized by reductions in their functional, cognitive and general clinical status, and present a significant deterioration in quality of life and high mortality rates. As a consequence, they require continuous attention from a wide variety of health professionals, rather than requiring attention in response to a specific problem. The costs of care that result from this continuous and frequent care represent an important part of the budgets of the health system.

Proper management and provision of care in the context of high-risk chronic patients involves focusing efforts on aligning health systems towards integrated care, and case management for high-risk patients in particular. The goal is to offer proactive, patient-centered care that is coordinated around select high-risk patients by assigning a referral physician or a small multidisciplinary team. For this, the individual needs of the patient are assessed and a care plan is developed accordingly.

The Navarra Health Service - Osasunbidea; SNS-O, implemented an integrated case management care program for high-risk chronic patients, in which the conditions to enrol into the program are as follows: Patients suffer at least three non-cancerous pathologies that include heart failure, dementia, ischemic heart disease, cerebrovascular disease, diabetes, chronic obstructive pulmonary disease, asthma, chronic kidney failure and cirrhosis, and patients belong to the top 5% of the risk pyramid.

A fundamental aspect of case management is to identify high-risk individuals based on a risk or probability of worsening their health status, a process known as risk stratification.

The first step in risk stratification is to estimate a risk score for each patient. The population is then segmented into different patient groups according to the resulting risk scores. Patients in each segment or group must have similar health care needs for the group to be useful for planning, and at the same time, each group must be different enough to warrant separate consideration. The key element of risk stratification procedures is that they must result in homogeneous groups of patients so that health programs can be designed and applied to all members of those groups.

But based on the available evidence, it is very possible that patients included in case management programs have heterogeneous needs. In particular, those patients at higher risk, at the apex of the risk pyramid, could benefit from a palliative care program rather than the case management program.

Objective

This thesis aims to answer the key question of whether the high-risk MCC population included in case management programs is heterogeneous in terms of risk.

For this purpose, a risk stratification is presented that determines whether and how many subpopulations of patients exist, as well as their particular characteristics. Subsequently, different options are presented to organize and plan care for each resulting subpopulation. This stratification will help to identify those subgroups of patients who do not benefit from their current care and to adapt the care strategies for them, directing the appropriate care to the appropriate patients. It also helps improve the efficiency of care for high-risk MCC patients. Finally, the survival patterns of the resulting patient subgroups were explored, in order to estimate the time to death for each type of patient. These data help plan end-of-life care for the population of interest.

One of the novelties presented in this thesis is that Artificial Intelligence (AI) methods were used to stratify the population into different subgroups. In particular, machine learning algorithms were used.

Main results

In general, patients have higher levels of risk as they age and their functional, nutritional and renal status decreases. Additionally, pressure ulcers are associated with higher risk scores, as well as clinical severity and high nursing care needs.

Three distinct patient subpopulations are identified among high-risk patients with ECM in the SNS-O; each one with its own characteristics and needs. Therefore, this thesis demonstrates that the population of patients with high-risk ECM included in case management programs is heterogeneous. Once the existence of subpopulations is demonstrated, organizational improvements in integrated care are suggested, supported by the evidence presented in this thesis.

It is suggested to continue with the current case management program for the resulting lower-risk subpopulation. It is also suggested to introduce a new case management program with an emphasis on home care for the intermediate-risk subpopulation that has limited functional capacity. Lastly, a home palliative care program is suggested for the very high-risk subpopulation.

Regarding the palliative care program for chronic non-cancer patients, one of the main barriers to its implementation is that determining the moment at which a patient with non-cancer MCC begins the final phase of life is complicated, due to the episodes of entry-re-entry of the disease. Health professionals have argued that, unlike cancer patients, the end of life in these patients is not predictable. The consequence is that palliative care is often delayed and most of the times it is not started.

This thesis demonstrates using survival models that estimates of time to death are in fact predictable in patients with non-cancer MCC, as a consequence of the identification of subpopulations of patients. As the high-risk population is heterogeneous, a survival pattern cannot be predicted for all patients simultaneously in the same model. However, if time to death is predicted in the new subgroups, given that if they are homogeneous, it is possible to make such predictions.

This has important implications since the results presented help to overcome obstacles in the implementation and organization of new non-cancer palliative care programs. Non-oncological MCC patients should no longer be excluded from these types of services, since it is now possible to estimate the time to death for each individual patient in the population of interest.

The early identification of the needs of patients helps to develop objective criteria for a correct and timely treatment of those patients in whom the terminal phase of their lives will occur in the near future. In the same way, it helps to plan health services.

Conclusion

This thesis aims to improve the health outcomes and care provided to high-risk MCC patients currently included in the SNS-O case management program. The risk stratification presented here achieves this purpose, identifying types of patients and helping to organize and plan the care that is better adapted to the needs of each one of them.

RESUMEN EJECUTIVO

Antecedentes

Los pacientes con múltiples enfermedades crónicas (MEC) se caracterizan por reducciones en su estado funcional, cognitivo y clínico general, además de presentar un importante deterioro de la calidad de vida y altas tasas de mortalidad. Como consecuencia, requieren atención continua de una amplia variedad de profesionales de la salud, y no atención en respuesta a un problema puntual. Los costes de atención que resultan de estos cuidados continuos y frecuentes representan una parte importante de los presupuestos del sistema de salud.

Una adecuada gestión y prestación de cuidados en el contexto de los pacientes crónicos de alto riesgo pasa por enfocar esfuerzos en el alineamiento de los sistemas de salud hacia la atención integral, y la gestión de casos para pacientes de alto riesgo en particular. El objetivo es ofrecer una atención proactiva y centrada en el paciente que se coordine en torno a pacientes seleccionados de alto riesgo mediante la asignación de un médico de referencia o un pequeño equipo multidisciplinario. Para ello se evalúan las necesidades individuales del paciente y se desarrolla un plan de atención en consecuencia.

El Servicio Navarro de Salud - Osasunbidea; SNS-O, implementó un programa de atención integral de gestión de casos para pacientes crónicos de alto riesgo, en el que las condiciones para entrar al programa son las siguientes: Los pacientes sufren al menos tres patologías no cancerosas que incluyen insuficiencia cardíaca, demencia, cardiopatía isquémica, enfermedad cerebrovascular, diabetes, enfermedad pulmonar obstructiva crónica, asma, insuficiencia renal crónica y cirrosis, y los pacientes pertenecen al 5% superior de la pirámide de riesgo.

Un aspecto fundamental en la gestión de casos es identificar a los individuos de alto riesgo en función de un riesgo o probabilidad de empeoramiento de su estado de salud, un proceso conocido como estratificación del riesgo.

El primer paso de una estratificación de riesgo es estimar una puntuación de riesgo para cada paciente. Después, la población se segmenta en diferentes grupos de pacientes de acuerdo con las puntuaciones de riesgo resultantes. Los pacientes de cada segmento o grupo deben tener necesidades de atención médica similares para que el grupo sea útil para la planificación y, al mismo tiempo, cada grupo debe ser lo suficientemente diferente para justificar una consideración separada.

El elemento clave de los procedimientos de estratificación del riesgo es que deben resultar en grupos homogéneos de pacientes para que los programas de salud puedan diseñarse y aplicarse a todos los integrantes de esos grupos.

Pero según la evidencia, es muy posible que los pacientes incluidos en programas de gestión de casos tengan necesidades heterogéneas. En particular, aquellos pacientes con mayor riesgo, en la cúspide de la pirámide de riesgo, podrían beneficiarse de un programa de cuidados paliativos en lugar del programa de gestión de casos.

Objetivo

Esta tesis tiene como objetivo responder a la pregunta clave de si la población de pacientes con MEC de alto riesgo incluida en los programas de gestión de casos es heterogénea en términos de riesgo.

Para ello, se presenta una estratificación de riesgo que determina si y cuántas subpoblaciones de pacientes existen, así como sus características particulares. Posteriormente, se presentan diferentes opciones para organizar y planificar los cuidados para cada subpoblación resultante. Esta estratificación ayudará a identificar aquellos subgrupos de pacientes que no se benefician de su atención actual y a adaptar las estrategias de atención para ellos, dirigiendo la atención adecuada a los pacientes adecuados. También ayuda a mejorar la eficiencia de los cuidados de los pacientes con MEC de alto riesgo. Finalmente, se exploraron los patrones de supervivencia de los subgrupos de pacientes resultantes, con el objeto de estimar el tiempo hasta la muerte para cada tipo de paciente. Estos datos ayudan a planificar los cuidados de final de vida para la población de interés.

Una de las novedades que se presentan en esta tesis es que se utilizaron métodos de Inteligencia Artificial (IA) para estratificar la población en diferentes subgrupos. En particular, se utilizaron algoritmos de aprendizaje automático (“machine learning”).

Resultados principales

En general, los pacientes tienen mayores niveles de riesgo a medida que envejecen y disminuye su estado funcional, nutricional y renal. Además, las úlceras por presión se asocian con puntuaciones de riesgo más altas, así como con la gravedad clínica y las altas necesidades de cuidados de enfermería.

Se identifican tres subpoblaciones de pacientes distintas entre los pacientes de alto riesgo con MEC en el SNS-O; cada una de ellas con sus características y necesidades particulares. Por lo tanto, esta tesis demuestra que la población de pacientes con MEC de alto riesgo incluida en los programas de gestión de casos es heterogénea. Una vez se demuestra la existencia de subpoblaciones, se sugieren mejoras organizativas en la atención integrada, respaldadas por la evidencia presentada en esta tesis.

Se sugiere seguir con el programa de gestión de casos actual para la subpoblación resultante de más bajo riesgo, introducir un nuevo programa de gestión de casos con énfasis en la atención a domicilio para la subpoblación de riesgo intermedio que tiene su capacidad funcional limitada, y por último se sugiere un programa de cuidados paliativos a domicilio para la subpoblación de muy alto riesgo.

En cuanto al programa de cuidados paliativos para pacientes crónicos no oncológicos, una de las principales barreras para su implementación es que determinar el momento en el que un paciente con MEC no oncológico comienza la fase final de la vida es complicado, debido a los episodios de entrada-reingreso de la enfermedad. Los profesionales de la salud han argumentado que, al contrario que los pacientes oncológicos, la fase final de la vida en estos pacientes no es predecible. La consecuencia es que los cuidados paliativos suelen retrasarse y la mayoría de las veces no se inician.

Esta tesis demuestra usando modelos de supervivencia que las estimaciones del tiempo hasta la muerte son de hecho predecibles en pacientes con MEC no oncológicos, como consecuencia de la identificación de subpoblaciones de pacientes. Como la población de alto riesgo es heterogénea, no puede predecirse un patrón de supervivencia para todos los pacientes. Sin embargo, si se predice el tiempo hasta la muerte en los nuevos subgrupos, dado que si son homogéneos, es posible hacer tales predicciones.

Esto tiene importantes implicaciones ya que los resultados presentados ayudan a superar los obstáculos en la implementación y organización de nuevos programas de cuidados paliativos no-oncológicos. Los pacientes con MEC no oncológicos ya no deberían ser apartados de este tipo de servicios, ya que ahora es posible estimar el tiempo hasta la muerte de cada paciente individual en la población de interés.

La identificación temprana de las necesidades de los pacientes ayuda a desarrollar criterios objetivos para un correcto y oportuno tratamiento de aquellos pacientes en los que ocurrirá la fase terminal de su vida en un futuro próximo. De igual manera, ayuda a planificar los servicios sanitarios.

Conclusión

Esta tesis busca mejorar los resultados de salud y la atención brindada a los pacientes con MEC de alto riesgo incluidos actualmente en el programa de gestión de casos del SNS-O. La estratificación del riesgo presentada aquí consigue este propósito, identificando tipos de pacientes y ayudando a una organización y planificación de la atención que se adapte mejor a las necesidades de cada uno de ellos.

GENERAL BACKGROUND

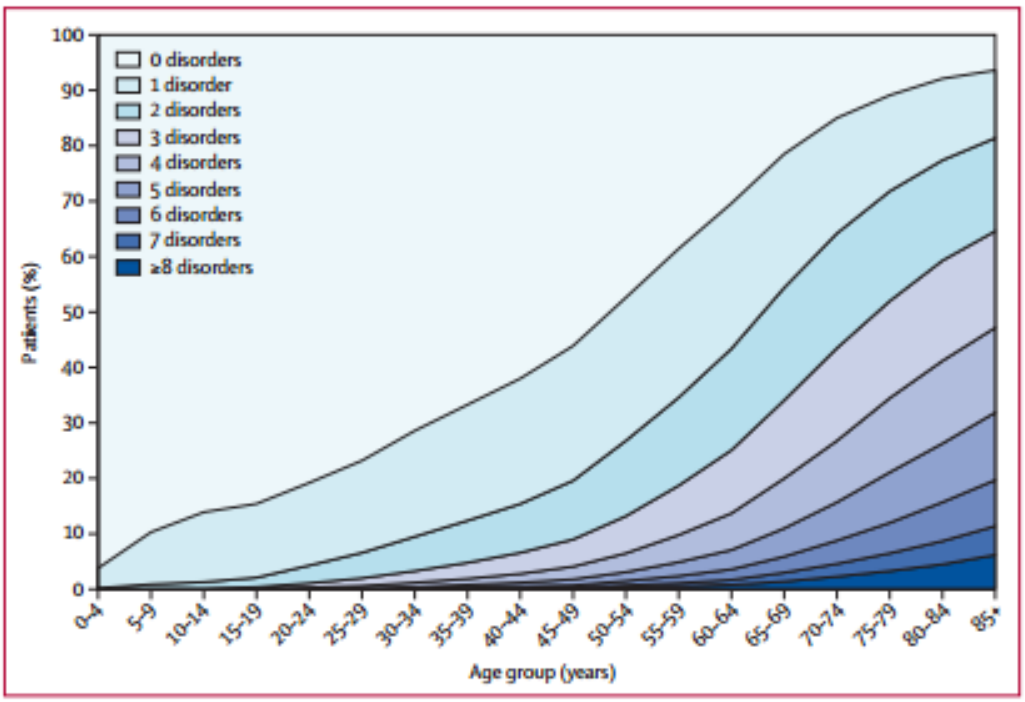
INTRODUCTION

The increase in life expectancy, together with the evolution of living and working conditions have produced an increase in the number of people that reach older ages. And as the world's population ages, individuals are significantly more prone to suffering chronic diseases [1, 2]. Consequently, the prevalence of chronic diseases is reaching alarming levels and non-communicable diseases are now the main causes of death in developed countries [3, 4].

In Europe, estimates of prevalence of chronic conditions indicate that more than 42% of the general population suffer one or more chronic conditions, and 23% of the population are diagnosed with more than one chronic illness at a time [5]. These patients with Multiple Chronic Conditions (MCC) are also known as multimorbid patients or polypathological patients.

The probability of suffering several chronic illnesses increases dramatically in old strata of the population, as seen in Figure 1[5]. The general MCC patient profile is characterized by reductions in functional, cognitive and overall clinical status, deterioration of quality of life, and high mortality rates. MCC patients also suffer frequent decompensations associated to entry-re-entry disease trajectories and require continuous attention from qualified professionals [3, 6-8].

Figure 1. Number of chronic disorders by age group [5].



The management of decompensations has been traditionally approached by healthcare organisations conceived to deal with acute care, which plan and administer treatment in response to punctual patient's needs. However, MCC patients have specific, complex care needs that are unlikely to be met in such settings. Unlike acute diseases that dominate healthcare delivery, MCC patients require care repeatedly from a wide variety of healthcare professionals that are not coordinated between them, making treatment harmonisation difficult and increasing the risk of adverse results.

When assistance is fragmented, meaning that there is a lack of continuity in treatment strategies, patients face an environment that is not conducive to improvements in health outcomes or quality of life. In addition to negative impact on health results, poor treatment coordination also leads to duplicative and inefficient care that increases healthcare utilisation rates.

As a result of the elevated number of consultations and treatments from different specialties, healthcare costs associated to chronic illness account for a high share of health system budgets [9]. There is an increase the pressure that health systems experience, because MMC patients are intensive in healthcare utilisation. Both presently and in coming decades, the approach towards MCC patients will be one of the greatest challenges facing governments and healthcare systems in terms of health outcomes and financially.

The MCC challenge is expected to become more prominent in the near future, as population pyramids in developed countries predict that the number of elderly patients will increase significantly. The 'baby boom' generation is approaching old age [10], and therefore the number of MCC patients will grow uncontrollably.

The organisational structure of acute healthcare systems is not prepared to deal with the deficiencies described above. Plans for an appropriate management and delivery of care in the context chronic illness and multimorbidity propose to focus efforts the realignment and redesign of systems towards integrated care [6, 8, 9, 11, 12].

INTEGRATED CARE

Integrated care is the new paradigm of organisation that has emerged as a response to the deficiencies of acute organisations. It emphasizes the coordination of multidisciplinary teams across all levels of care, seeking better outcomes, quality of life and efficiency in healthcare utilisation. The objective is to provide comprehensive, proactive and patient-centred care rather than acting reactively. In other words, health systems should transit to a proactive, patient-centred holistic approach rather than maintaining a reacting, disease-focused perspective. In this way, professionals can deal with chronic patients' multidimensional needs while keeping follow-up [13].

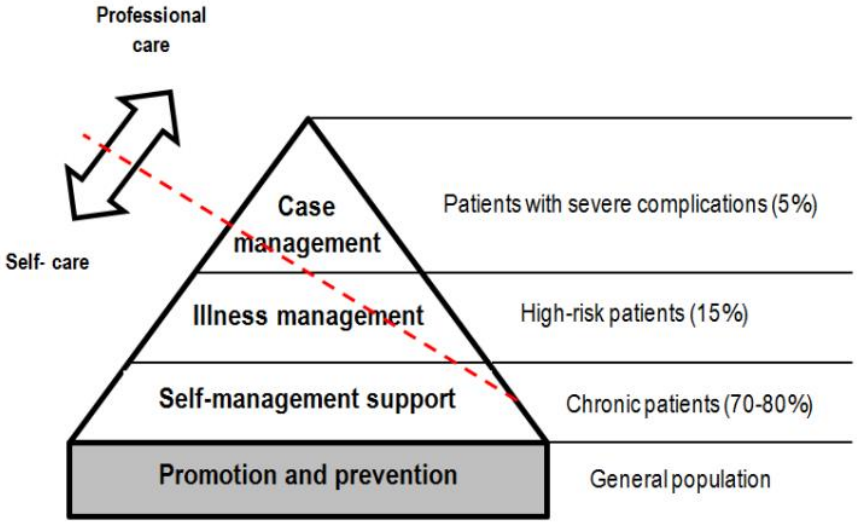
There is a variety of ways in which integrated care can be structured, depending on the severity level or risk profile of the target population for which the programme is addressed. Programmes can be classified into three possible levels of integration. These are population-based models (macro level), group-based models (meso level) and individual-based models (micro level) [14-16]:

- Population-based models

These models deliver integrated care services to all patients, independently of the risk stratum to which they belong. However different individuals require different type of services depending on their risk profile. Therefore, population-based models rely on risk stratification for the identification of homogeneous groups of patients in order to adapt care to their needs.

All patients receive services centred in prevention and promotion, together with self-management support, and as risk increases patients also receive Disease Management (meso level integrated care) and Case Management (micro level integrated care) services. Other characteristic elements of population-based models include the creation of new professional groups that specialise in different patient risk profiles, the use of information systems or the guiding in the transition of patients across different risk strata. Examples of population-based integrated care programmes include Kaiser Permanente in the United States, and its well-known stratification model represented in their risk pyramid (Figure 2) [13, 16, 17].

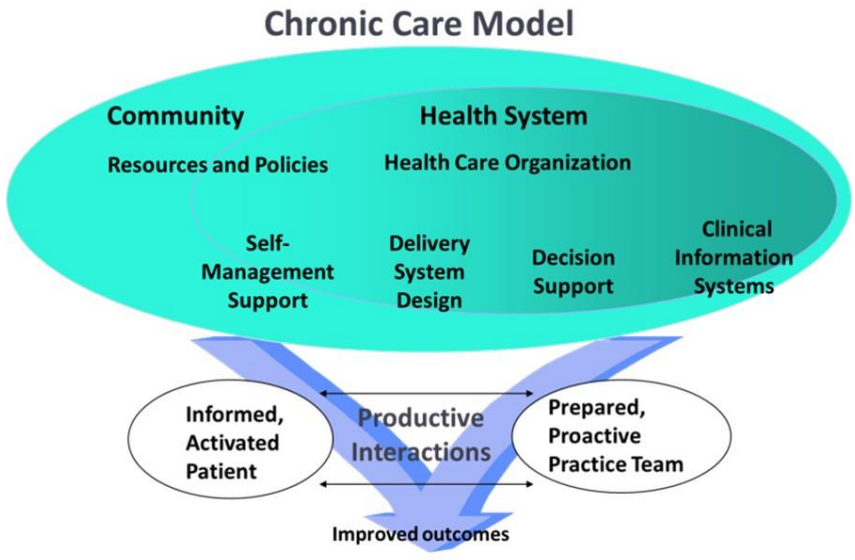
Figure 2. Population-based integrated care models. The Kaiser pyramid [13]



- Group-based models

These models focus on specific subpopulations that have been identified following the application of risk stratification processes to population-based models, and are known as Disease Management programmes. The identification of subgroups allows health systems to design specific programmes for each of them, framing organisational efforts around groups of chronic patients with the same risk profile, chronic condition or combination of chronic illnesses. The most well-known and applied programme of this nature is the Chronic Care Model (CCM), developed by Wagner and colleagues (Figure 3) [11, 12, 18].

Figure 3. Group-based integrated care models. The Chronic Care Model [12, 18]



The CCM was developed as a result of a systematic literature review that identified intervention components that improve chronic patients' health outcomes, quality of life and system efficiency [11, 12, 18]. The model does not provide a set of specific interventions to be implemented, but rather a generic framework that can be applied across all organisations and settings.

Areas in which organisational efforts focus include the use of guidelines and protocols, practice re-design to meet changing patient needs, patient self-management and education, professional expertise and decision support, and information systems. The model also includes Case Management features as patient risk increases.

The key to success are the productive interactions between informed, proactive patients and prepared, expert practice teams that take a holistic approach to patient care. In cooperation, both patients and professionals take an active part in the design of care strategies that lead to improved outcomes and healthcare utilisation efficiency. Further modifications of the model include the Expanded Chronic Model, which makes emphasis in prevention and promotion [7], or the Innovative Care for Chronic Conditions Framework [19].

- Individual-based models

This level of integration focuses on coordinating programmes around selected high-risk patients and their carers. For this purpose, it is essential to identify those individuals by means of risk stratification techniques. Once high-risk individuals have been identified, Case Management can be provided to those patients.

Case management refers to the planning and the coordination around specific cases through the assignment of a reference physician or a small multidisciplinary team from different professional backgrounds, not necessarily related to healthcare [13].

Teams assess the individual needs of each patient, develop a care plan accordingly and coordinate treatment delivery. Patients are monitored in periodic re-assessments and treatment plans are adjusted if needs change. Information systems such as electronic medical records play a central role in patient follow-up and in the supply of data for risk stratification purposes.

RISK

Given that integrated care advocates for a patient-centred approach, healthcare services should address the specific needs and characteristics of each individual patient to achieve optimal health. However, it is virtually impossible to develop healthcare models or treatment strategies that satisfy the particular needs of every patient in a personalised way. Instead, healthcare organisations design care strategies that can be applied to groups of patients with largely similar level of risk and clinical characteristics [17]. This is the case for integrated care models presented above, which target different patient types.

The identification of these groups of patients is known as risk stratification. Risk stratification processes group patients in terms of their risk, or probability, of worsening in their health status. Risks are estimated for each patient using both statistical models and judgements from clinicians [9].

Statistical processes in this context usually consist of predictive models that allow determining the probability of the occurrence of a particular outcome in a given time horizon, such as dying, being a high cost patient, or any other outcome that may be of interest in the design of care strategies for each of those groups or patient types. This estimation requires to establish relationships between clinical data that are usually registered in electronic health records.

Once a risk score has been estimated for every patient, the population can be segmented into different patient groups according to the individual patient risk scores, or equivalently, the estimated probability of suffering the defined adverse event. Patients in each segment or group should have similar health care needs to make the group useful for planning, and at the same time each group must be different enough to justify separate consideration [20]. Segmentation processes are in the infancy and there are not many defined methods [21], yet artificial intelligence techniques are increasingly being used [22].

The result are homogeneous groups of patients with similar clinical needs, and level of morbidity. Stratifying the population of interest into similar subgroups allows the delivery of specific programmes tailored to them. Different groups or strata will require more intense case and greater healthcare utilisation. For example, very high risk patients

will require case management services, while low risk patients may only need self-management support.

It can also help to identify a target population, and to identify key characteristics inherent to that population that can guide the design of treatment strategies to fit the specific needs of each patient group. Exploring the characteristics of each of patient type and understanding their clinical needs helps to set eligibility criteria for each programme.

Tailored programmes improve outcomes while reducing or preventing the use of costly services; maximize efficient use of health care resources [21]. On the contrary, the lack of risk stratification results in the inclusion of heterogeneous populations in programmes that do not fit them, resulting in worse outcomes and resource wastage.

EXPERIENCE IN THE NAVARRA HEALTHCARE SERVICE

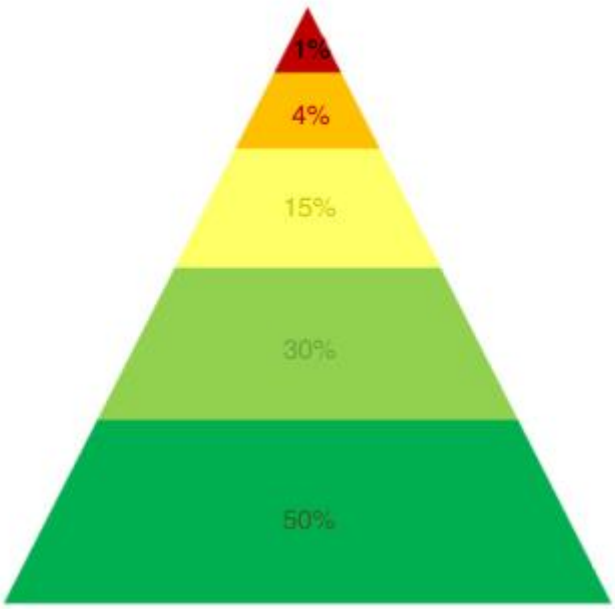
The Spanish region of Navarra and its public health system, the Navarre Healthcare Service (Servicio Navarro de Salud – Osasunbidea; SNS-O), covers the majority of the population of Navarra, around 650,000 inhabitants. At present, chronic diseases in general account for more than 70% of consultations in primary care, more than 60% of hospital admissions and more than 25% of pharmaceutical prescriptions issued in the SNS-O; in total more than 10,800,000 prescriptions [23]. The economic impact is outstanding and affects the sustainability of the entire public healthcare system.

In recent years, the SNS-O has been making an important effort to try to provide integrated care for chronic patients in general, with an additional case management program for severe MCC patients. The program is known as the “Integrated Care Strategy for Chronic and Multiple Chronic Conditions Patients” (Estrategia Navarra de Atención Integrada a Pacientes Crónicos y Pluripatológicos) and is intended to provide appropriate care for these patients, improving health outcomes and at the same time tackling possible problems of economic sustainability.

A strategic line of the integrated care program is the stratification of the whole chronic population, without distinguishing which or how many illnesses were diagnosed in each patient. Different levels of severity were sought to be identified, in the style of the population-based integrated care models such as the Kaiser pyramid explained above.

Since 2014, Navarra has used the GMA (Grupos de Morbilidad Ajustados; Adjusted Morbidity Groups) [24] as a risk stratification tool to divide the entire chronic population in the region. The GMA is widely used in Spain and is similar to the Clinical Risk Groups (CRGs), one of the best-known stratification tools worldwide. The GMA divides the population into 5 subgroups or levels of severity using percentiles, as shown in Figure 4. Like CRGs, its main purpose is to identify groups of patients with a similar level of healthcare utilisation.

Figure 4. GMA levels of severity [24]



In this way, patients receive different types of interventions depending on their level of severity. People with greater severity or complexity benefit especially from a case management care model, without forgetting the responsibility of self-care that is possible at each level of severity. At the same time, the adaptation of care to the needs of each patient type allows optimisation of care and a better allocation of resources.

Three different levels of intervention are distinguished according to their level of severity measured through the GMA and the characteristics of the patients. These levels follow the structure of integrated care that includes population, disease and case management models:

- Level 1: The majority of chronic patients diagnosed at this level correspond to patients of low complexity and easy control. They receive support for their self-management and self-care.
- Level 2: Patients diagnosed with a high level of risk, but of less complexity (although in an advanced disease stage). They receive disease management care that combines self-management and professional care as in the Chronic Care Model.
- Level 3: Patients diagnosed with a high level of risk and high complexity in their care and treatment, and very frequently MCC patients. It is necessary to carry out a comprehensive case management approach which consists fundamentally of professional care. In these cases, a very exhaustive follow-up is needed to control the patient.

Those included in level 3 of care, which correspond to the case management integrated care program, must be classified into levels 4 or 5 of the GMA Scale. This means that they are very severe patients that represent the top 5% risks of the population pyramid. Note that the whole population is stratified and assigned to a level of care. However, very high risk MCC patients have a special case management program for them, due to the elevated costs they implicate and the complexity of their needs.

MULTIPLE CHRONIC CONDITIONS CASE MANAGEMENT PROGRAMME

The SNS-O determined in 2011 that MCC patients generated around 100,000 primary care consultations and a 100-bed hospital would be needed to attend them, regardless of their severity level. It is common for MCC patients to go more than once a week to the primary care centre, and they have an average of 12 visits to the emergency services, 6 specialist visits and 18 days of hospital admission per year. These numbers have increased in the last decade, and the challenge is particularly great among high-risk severe MCC patients. The average cost of these individuals is 33 times higher than that of a patient who does not suffer from any chronic disease [25]. Despite the fact that there are only a few hundred severe MCC patients, they represent a large share of healthcare spending, putting great pressure on the budget.

In addition, high-risk MCC patients also present very high mortality rates, low functional status and adverse clinical outcomes. Patient care in these cases is very complex and difficult to standardize, and therefore an integrated care plan is in great need. For this reason, and given that it is directed to very high-risk patients, the SNS-O implemented a case management model for this particular population.

The conditions to qualify for program enrollment were as follows: Patients suffered at least three selected noncancer pathologies, including heart failure, dementia, ischemic heart disease, cerebrovascular disease, diabetes, chronic obstructive pulmonary disease, asthma, chronic renal failure and cirrhosis, and patients belonged to the top 5% of the risk pyramid according to GMA risk stratification tool.

It is important to point out that this case management programme does not consider cancer patients. In other words, only non-oncologic patients are enrolled. This can be explained as cancer patients have different disease progression patterns in comparison to the rest of chronic illnesses [26]. Cancer patients will experience different outcomes and will react in their own way to treatments. For example, they will have different life expectancies in comparison to non-oncological patients. For this reason, cancer their own particular programmes of care in Navarra and are not treated under the MCC case management integrated care strategy.

At patient enrolment into the program, a comprehensive assessment of the patient's clinical condition is carried out. In this assessment, professionally rated variables including clinical severity, nursing needs and social needs are recorded. Data is registered in the electronic medical records of primary care and specialized care. In addition, sociodemographic data is also available, as well as data regarding functional status, nutritional status, renal deterioration status, the presence of pressure skin ulcers, the number of prescriptions, prevalence and number of coexisting selected illnesses and the GMA risk score of each high-risk MCC patient.

The services that are included as part of the case management strategy include the assignment of a specialist of reference to follow the case of each patient, with joint follow-up with primary care. In this follow-up, a nutritional risk assessment is made, a medication review in the case of polymedicated patients (more than 8 prescriptions for chronic illnesses), patients can telephone a hospital liaison nurse, and there is also a day hospital. It also seeks greater agility in the care circuits, giving the possibility of direct

admission without having to go through the emergency room, discharge is planned early, and self-care is promoted from the moment the patient is admitted. Once the patients are discharged, continuity in the follow-up is maintained via primary care or nursing during one day.

ARE HIGH-RISK MCC PATIENTS' NEEDS HETEROGENEOUS?

The authorities from the SNS-O performed an assessment of the case management programme after its implementation (unpublished), in which they analysed the evolution of costs and survival rates among two different groups of patients. Both groups were high-risk MCC patient with similar characteristics, but one of them was enrolled into the case management programme whereas the other group received the standard of care.

One of these analyses made a distinction between those patients who lived and those who died. Results showed that in the living patients, the intervention led to a lower increase in costs. In other words, both the case management and standard of care groups had an increase in costs but the increase was significantly lower in the case management group. This effect was due to the number of averted emergency care episodes and also the number of averted hospital admissions. However, the dead patients in both groups showed a similar increase in costs. Importantly, there were no differences in mortality rates between the standard of care and case management groups.

This means that the intervention had an effect on those patients that managed to survive, but it did not on those patients that died. A similar proportion of patients still died in comparison to standard of care, despite being in the case management group. That is to say, the programme was successful only for a subgroup of the high-risk MCC patients.

The key element of risk stratification procedures is that they should result in homogeneous groups of patients so that the healthcare programmes can be designed and applied to everyone in that groups, based on the premise that everyone has similar characteristics and needs. But based on the evidence of the case management programme for high-risk MCC patients in the SNS-O, it may be very possible that the patients included in such programme have heterogeneous needs.

In particular, those patients with the greatest risk, at the apex of the risk pyramid, may benefit from palliative care instead of the case management programme. The new integrated care models should extend from the moment of diagnoses until the end of

life, covering all the ill life of the patient. It is therefore necessary to guarantee good care and a quality of life until the moment of death. Care models should then include palliative care.

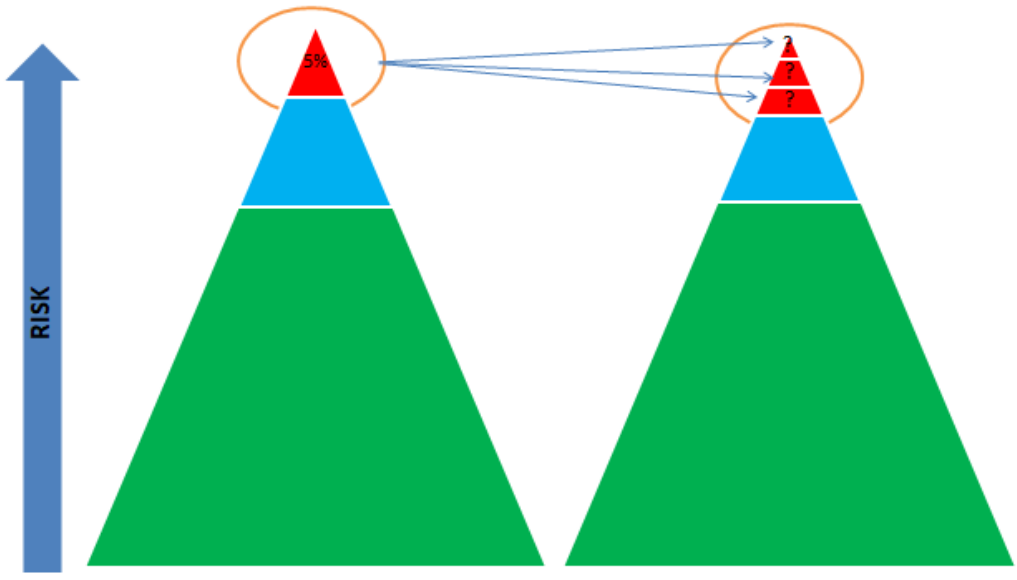
The SNS-O currently has a palliative care scheme in place, which covers cancer patients in its majority. However, there are plans to extend this kind of care to non-cancer MCC patients to address this lack of necessary treatment. It is intended to improve access to care and to complete the deployment of palliative care to all kinds of patients, no matter of their diagnoses. Non-oncologic patients are equally deserving and severe to consider them for palliative care.

However, one aspect needs to be clarified. The criteria to include cancer patients into palliative care are well defined, accepted and applied. Yet it is not clear what type of patients affected by chronic, non-oncologic diseases will really benefit from these new care modalities [25].

GENERAL THESIS OBJECTIVES

This thesis aims to answer the key question of whether the high-risk MCC patient population included in case management programmes is heterogeneous in terms of risk. For this purpose, this study produces a risk stratification that will identify if and how many patient subpopulations exist, as well as their particular characteristics (Figure 5).

Figure 5. Risk stratification thesis objectives.



This stratification will help to identify those patient subgroups that do not benefit from their current care and to adapt care strategies for them. Especially, identifying those patients that may benefit from palliative care is essential for the general strategic lines of the case management programme. The purpose is to target appropriate care for the appropriate patients.

A number of complementary objectives are proposed in line with this main objective:

- To develop a risk score predictive model, which can be used to assign a risk score to each particular patient in the case management programme.
- To identify predictive indicators of risk that can help the development of objective criteria for a correct and timely identification of different types of high risk patients. Especially those very high risk at the apex of the risk pyramid that might be in need of end-of-life care.
- To provide with instruments to apply the risk score predictive model in clinical settings. This instrument should be easy to use for healthcare professionals to calculate individual risk scores.
- To stratify the population according to the estimated patient risk score for each MCC patient.
- Once the patient subgroups are identified, to determine the defining characteristics and needs of each type of patient.
- Given the particular characteristics and needs of each patient subgroup, propose different care strategies for each of them. These care strategies should adapt to the needs of each patient type.
- Since one of the key points of the case management programme is a potential extension of palliative care for non-oncologic MCC patients, one of the patient subgroups shall be probable candidates for end-of-life care. Therefore an additional aim is to estimate survival time to death in each of the non-oncologic patient subgroups that result from the risk stratification procedures. This will enhance end-of-life care in high-risk MCC patients and it will ease early identification of potential palliative care users.
- To determine whether survival in non-cancer high-risk MCC patients is predictable and to find prognostic variables that predict time to death.

GENERAL THESIS METHODS

The thesis is divided into three main chapters. Each of them corresponds to a key element of risk stratification procedures. However, all three chapters use the same database, and therefore have common variables and participants. The database is anonymised and i aggregates demographic data together with other data from several sources, including primary care and hospital electronic clinical , hospital nursing history, and electronic prescriptions data. All thesis calculations were performed using the STATA software.

MCC patients that are in the database belong to the top 5% of the case management programme's risk pyramid according to the Adjusted Morbidity Groups (GMA) stratification tool, and suffer at least three of the following non-cancer pathologies: Heart Failure, Dementia, Ischemic Heart Disease, Cerebrovascular Disease, Diabetes, Chronic Obstructive Pulmonary Disease, Asthma, Chronic Renal Failure or Cirrhosis. A total of 885 patients are considered, representing all high-risk non-oncological MCC patients of the region of Navarra.

At enrolment, patients are subject to a comprehensive assessment of their situation. A set of baseline variables are collected at this point in time, including age, sex, functional status (Barthel scale Lawton & Brody scale, and degree of dependency), nutritional status (serum albumin level), renal deterioration status (creatinine level, albumin/creatinine index), presence of pressure skin ulcers, and global status, which is a variable indicating patient severity. Global status score is produced by the case management team responsible for each patient (doctors, nurses and social services workers) based on expert professional criteria. Other professional-rated scales informing of particular areas of interest such as clinical severity, nursing needs or social needs are available. Prevalence of selected non-cancer illnesses is also recorded. The date of inclusion into the programme, as well as the date of death are available. The individual GMA score of each patient is also recorded.

Not all variables were collected for all patients, and therefore the database contains missing values. To fill in missing values Multiple Imputation techniques were used with the *chained equations* method [27]. This is done so that all the information for all the patients can be used in estimations throughout the thesis, instead of excluding those

patients with incomplete information. Therefore, we do not miss any data and we avoid using a partial and skewed sample of the high-risk MCC patient population.

If Multiple Imputation would not done, the sample to be used in calculations would not be representative of the population, and thus it is better to use imputed values for those missing values and use the whole patient population of the region. As a consequence, results are not biased. Additional information can be found in Supplementary Material 1.

Chapter 1

The first part of a risk stratification consists in assigning a risk score to each patient that is to be stratified based on that particular score. The outcome for this risk score estimation was 1-year mortality from enrollment in the case management programme. As a consequence, a logistic regression model was fitted using data from the initial comprehensive assessment that was completed upon inclusion. The reason why mortality was used as the outcome was that our population of interest consisted of patients with different chronic illness combinations. Therefore, disease-specific outcomes were not appropriate. Additionally, it was sought to stratify the population in terms of a clinical variable, rather than risk scores based on cost or utilisation.

The relationship off all variables with 1-year mortality outcomes was first assessed in univariate analyses. Those variables that were significant at a p-value level of $p \leq 0.25$ were then fitted into a multivariate model using stepwise methods. Discriminative ability of the model was measured using the Receiver Operating Characteristic (ROC) curve and its associated Area Under the Curve (AUC). The Hosmer-Lemeshow test and calibration plots were used to assess goodness of fit. The model was validated internally using k-fold Cross-Validation with 10 random Bootstrap Validation. Results are expressed in terms of odds ratios (OR), yet regression coefficients are also available at Supplementary Materials 2.

Finally, we provide with tools to apply the risk score predictive model in clinical practices. To do so, a *nomogram* was constructed in STATA using the 'nomolog' command [28]. A nomolog is a graphical calculation tool that consists on a set of interconnected lines that allow the calculation of individual patient risk scores in question of minutes.

Chapter 2

Once a risk score is available for each patient, the whole MCC population was categorized into different subgroups according to their estimated risk scores. Subgroups represent mutually exclusive risk strata, and machine learning algorithms were used to build them. These techniques are generally called cluster analyses and group individuals who have similar risk scores into subgroups or clusters.

A particular type of algorithms were used, which were hierarchical algorithms and the Ward's linkage hierarchical algorithm with the squared Euclidian distance in particular. With this method, observations merge with other observations that are similar to them in terms of distance. Those observations that are closest to them, or equivalently those that have the most similar risk scores, are merged into a group. This process continues until the optimal number of clusters is reached. The optimal number of clusters was determined using the Duda/Hart stopping rule (Supplementary Material 3) and visually through a dendrogram. A dendrogram is a graphical way of illustrating the grouping process and is very informative in terms of deciding the optimal number of clusters.

The results were validated using Silhouette scores [29], which tell whether observations have been correctly assigned to the group with the most similar risk scores. To evaluate the stability of the results, the full sample was randomly divided into four equally sized subsamples, each containing 25% of the observations, and the algorithm was run again on each subsample as if it was a k-fold cross validation with four folds. Other algorithms such as the K-means algorithm were also used to test whether results were similar.

The characteristics of each subgroup were compared to test if there were significant differences between them. This was done using one-way ANOVA tests in the case of continuous variables, Kruskal-Wallis tests for categorical variables or if the assumptions for ANOVA did not hold, and $\chi 2$ tests for binary variables. Post-hoc pairwise analyses were completed to identify which cluster was different from the remaining clusters. Multiple one-way ANOVA comparisons with Bonferroni corrections, Mann-Whitney U tests and Fisher's exact tests were used if ANOVA, Kruskal-Wallis or $\chi 2$ tests were used, respectively.

Chapter 3

Since one of the key objectives of the thesis was to identify potential users for an extended palliative care programme for non-oncological MCC patients, it is crucial to know when patients will be needing these services. For this purpose, survival analyses were performed to estimate time to death for each patient subgroup.

For each patient, time to death since enrolment into the programme, or time to study end if patients were censored were computed. Kaplan-Meier analyses were performed to estimate survival time for all three high-risk patient subgroups. The logrank test was used to determine whether the survival curves were significantly different from each other. Cox proportional hazard regression analyses were performed to assess the relationship between patient characteristics and survival. The proportional hazards (PH) assumption of the model was tested evaluating the Schoenfeld residuals and visually by comparing the baseline survival function for each of the patient subgroups (Supplementary Material 4). Further parametric models were fitted to extrapolate survival. These models allow predicting survival beyond study period limits.

REFERENCES

1. Hajat, C. and E. Stein, The global burden of multiple chronic conditions: A narrative review. Prev Med Rep, 2018. 12: p. 284-293.
2. Navickas, R., et al., Multimorbidity: What do we know? What should we do? J Comorb, 2016. 6(1): p. 4-11.
3. Ham, C., The ten characteristics of the high-performing chronic care system. Health Econ Policy Law, 2010. 5(Pt 1): p. 71-90.
4. Wang, H., et al., Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: a systematic analysis for the Global Burden of Disease Study 2015. The Lancet, 2016. 388(10053): p. 1459-1544.
5. Barnett, K., et al., Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. Lancet, 2012. 380(9836): p. 37-43.
6. Fernández Moyano, A., et al., Care models for polypathological patients. Rev Clin Esp, 2017. 217(6): p. 351-358.
7. Barr, V.J., et al., The expanded Chronic Care Model: an integration of concepts and strategies from population health promotion and the Chronic Care Model. Hosp Q, 2003. 7(1): p. 73-82.
8. Glasgow, R.E., C.T. Orleans, and E.H. Wagner, Does the chronic care model serve also as a template for improving prevention? Milbank Q, 2001. 79(4): p. 579-612, iv-v.
9. Orueta, J.F., et al., FINGER (Forming and Identifying New Groups of Expected Risks): developing and validating a new predictive model to identify patients with high healthcare cost and at risk of admission. BMJ Open, 2018. 8(5): p. e019830.
10. Lynn, J., et al., Using population segmentation to provide better health care for all: the "Bridges to Health" model. Milbank Q, 2007. 85(2): p. 185-208; discussion 209-12.
11. Wagner, E.H., B.T. Austin, and M. Von Korff, Organizing care for patients with chronic illness. Milbank Q, 1996. 74(4): p. 511-44.
12. Wagner, E.H., Chronic disease management: what will it take to improve care for chronic illness? Eff Clin Pract, 1998. 1(1): p. 2-4.
13. A., S., Integrated care models: an overview. World Health Organisation. World Health Organization, 2016.
14. Valentijn, P.P., et al., Understanding integrated care: a comprehensive conceptual framework based on the integrative functions of primary care. Int J Integr Care, 2013. 13: p. e010.
15. Leutz, W.N., Five laws for integrating medical and social services: lessons from the United States and the United Kingdom. Milbank Q, 1999. 77(1): p. 77-110, iv-v.
16. Ham, C.C., N., Integrated care. What is it? Does it work? What does it mean for the NHS? The KIng's Fund, 2011.
17. Vuik, S.I., E.K. Mayer, and A. Darzi, Patient Segmentation Analysis Offers Significant Benefits For Integrated Care And Support. Health Aff (Millwood), 2016. 35(5): p. 769-75.
18. Wagner, E.H., Managed care and chronic illness: health services research needs. Health Serv Res, 1997. 32(5): p. 702-14.
19. Organization, W.H., Innovative Care for Chronic Conditions. Global Report. 2002.

20. Lynn, J. and D. Adamson, Living Well at the End of Life: Adapting Health Care to Serious Chronic Illness in Old Age. 2003.
21. Jean-Baptiste, D., O'maley, A., Shah T.,, Population Segmentation and Targeting of Health Care Resources: Findings from a Literature Review. Mathematica Policy Research, 2017. Working Paper 58.
22. Vuik, S.I., E. Mayer, and A. Darzi, Enhancing risk stratification for use in integrated care: a cluster analysis of high-risk patients in a retrospective cohort study. BMJ Open, 2016. 6(12): p. e012903.
23. Ibarrola Guillén C., S.B.A., Iribarren Udobro I., et Al., Estrategia Navarra de atención integrada a pacientes crónicos y pluripatológicos. 2013.
24. Monterde, D., E. Vela, and M. Clèries, [Adjusted morbidity groups: A new multiple morbidity measurement of use in Primary Care]. Aten Primaria, 2016. 48(10): p. 674-682.
25. Ibarrola Guillén C., L., J., Astiz, J., et Al., Evaluación Estrategia Navarra de Atención Integrada a pacientes pluripatológicos a los dos años de implantación. 2018.
26. Lenzi, J., et al., Burden of multimorbidity in relation to age, gender and immigrant status: a cross-sectional study based on administrative data. BMJ Open, 2016. 6(12): p. e012812.
27. White, I.R., P. Royston, and A.M. Wood, Multiple imputation using chained equations: Issues and guidance for practice. Stat Med, 2011. 30(4): p. 377-99.
28. Zlotnik, A. and V. Abraira, A General-purpose Nomogram Generator for Predictive Logistic Regression Models. The Stata Journal, 2015. 15(2): p. 537-546.
29. Rousseeuw, P.J., Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics, 1987. 20: p. 53-65.

CHAPTER 1

ESTIMATING RISK SCORES FOR HIGH-RISK MULTIPLE CHRONIC CONDITIONS PATIENTS

Published as:

"Multisystem chronic illness prognostication in non-oncologic integrated care"

at the BMJ British Medical Journal Supportive and Palliative Care

Citation:

Bretos-Azcona PE, Ibarrola Guillén C, Sánchez-Iriso E, et Al. Multisystem chronic illness prognostication in non-oncologic integrated care. BMJ Support Palliat Care. 2020 Jun 24:bmjspcare-2019-002055. doi: 10.1136/bmjspcare-2019-002055.

Presented at the conferences:

XXXIX Jornadas de la Asociación de Economía de la Salud (12-14 June 2019); Albacete, Spain

ABSTRACT

Objectives: To develop a mortality predictive model for a correct identification of non-cancer multiple chronic conditions patients that would benefit from palliative care, recognise predictive indicators of death, and provide with tools for individual risk score calculation.

Design: Retrospective observational study with multivariate logistic regression models.

Participants: All high-risk multiple chronic conditions patients incorporated into an Integrated Care strategy that fulfil two conditions: 1) They belong to the top 5% of the programme's risk pyramid according to the Adjusted Morbidity Groups (GMA) stratification tool and 2) They suffer simultaneously at least three selected chronic non-cancer pathologies (n=591).

Main outcome measure: 1-year mortality since patient inclusion in the programme.

Results: Among study participants, 201 (34%) died within the 1-year follow-up. Variables found to be independently associated to 1-year mortality were the Barthel Scale ($p<0.001$), creatinine value ($p=0.032$), existence of pressure ulcers ($p=0.029$) and patient global status ($p<0.001$). The Area Under the Curve (AUC) for our model was 0.751, which was validated using Bootstrapping (AUC= 0.751) and k-fold Cross-Validation (10 folds; AUC= 0.744). The Hosmer-Lemeshow test ($p=0.761$) showed good calibration.

Conclusions: This study develops and validates a mortality prediction model that will guide transitions of care to non-cancer palliative care services. The model determines prognostic indicators of death and provides tools for the estimation of individual death risk scores for each patient. We present a nomogram, a graphical risk calculation instrument that favours a practical and easy use of the model within clinical practices.

INTRODUCTION

Background

Multiple chronic conditions patients have become a growing concern for clinicians and health system administrators [1]. At present, approximately 50% of all individuals in the general population suffer at least one chronic disorder, and around one quarter suffer several chronic disorders at the same time [2].

This typology of patients is more common in elderly strata of the population [3], among which functional loss and frailty are high. The number of people reaching old age is growing, and consequently the prevalence of patients with multiple chronic diseases is increasing. [4].

As a result of the increase of prevalence of multimorbidity, concerns over health systems' sustainability are arising. Multiple chronic conditions patients use healthcare resources in an intensive and frequent way, and despite representing a small number of patients they account for the majority of the budget in some organisations [5, 6].

Efforts to contain cost and provide appropriate care focus on integrated care models that aim to treat multiple pathologies in a uniform and coordinated manner, improving patient quality of life and encouraging less use of resources. However, not all patients included in integrated care models benefit from them. Very severe patients, which belong to the apex of the Kaiser pyramid, have very high mortality rates and do not achieve improvements neither in their quality of life nor in their clinical progression despite high resource utilisation [7, 8].

These patients with advanced disease could benefit from programs focused on end-of-life care instead, focusing on improving their quality of life and putting aside ineffective and expensive treatments [9, 10].

Despite the fact that the number of deaths attributable to chronic non-cancer diseases exceeds those of cancer [11] and the potential benefits that palliative care could provide, this type of care has traditionally been reserved for cancer patients. There is evidence for this unmet need, with very low rates of non-oncological patients included in palliative care programmes. In Spain, France or Germany chronic patients in palliative care account for less than 10% of utilisation of these services, and at a European level, this figure drops to 6% [12].

Among the obstacles that are causing the delay in the implementation of these specific palliative care models, one in particular stands out. Determining exactly the moment in which the deterioration of a patient gives way to the terminal phase of life is complicated, especially in patients with more than one advanced-stage disease with "entry-re-entry" trajectories [9, 13].

The identification of predictive indicators of death can help the development of objective criteria for a correct and timely identification of those patients in whom the terminal phase of their life will happen in the near future. Obstacles presented by health professionals could be overcome, providing tools to avoid further delays in transitions towards palliative care for multiple chronic conditions patients.

Organisations like Medicaid and Medicare require a prognosis of ≤6 months to be eligible for hospice benefits [13]. However, predictive models that estimate 6-month mortality in non-cancer patients have shown a lack of ability to discriminate patients accurately. Frequent exacerbations and entry-re-entry trajectories make short-run estimation unpredictable in nature and consequently less accurate [9, 14].

Using longer time frames, such as 1-year mortality, is therefore more appropriate. Additional benefits arise from this approach, since it allows early identification of patients, rather than identifying them in their latest stages of life when the terminal decline has begun. This allows reorientation of clinical strategies towards quality of life improvements and reduction of patient distress, instead of continuing with aggressive treatments with no curative effects [15]. Moreover, this reduction in ineffective treatments and avoidable hospitalizations generates cost savings and greater efficiency to the system.

This study develops a 1-year mortality predictive model for its use as a risk assessment tool that can guide the correct identification of patients that are near to their terminal phase of their lives.

Objectives

Three objectives have been set in this study. The first one is to develop a mortality predictive model, the second objective is to identify predictors of death and the third objective is to provide with a nomogram to healthcare professionals to calculate individual risk scores.

METHODS

Data and Participants

The context of this study is the integrated care program for multiple chronic conditions patients of the Spanish region of Navarre (640.000 inhabitants), implemented in 2016. Data was obtained from the specific integrated care programme database, which is anonymised. This database is population-based, and aggregates demographic data together with other data from several sources, including primary care and hospital electronic clinical histories, hospital nursing history, intra and extra hospital electronic prescriptions, as well as degree of dependency recognition according to Spanish Law [16].

The study population consists of all high-risk multiple chronic conditions patients from the region incorporated into the integrated care strategy from April 2016 to August 2018 that completed at least a one year follow-up from enrolment, or died before follow-up completion. Those patients for which one year follow-up was not completed because the study period came to an end and had not died were excluded from analysis to avoid censoring problems

Programme conditions to qualify for enrolment are the following: 1) They belong to the top 5% of the programme's risk pyramid according to the Adjusted Morbidity Groups (GMA) stratification tool [17, 18, 19], which is similar to other tools such as Clinical Risk Groups (CRGs) and 2) they suffer at least three of the following non-cancer pathologies: Heart Failure, Dementia, Ischemic Heart Disease, Cerebrovascular Disease, Diabetes, Chronic Obstructive Pulmonary Disease, Asthma, Chronic Renal Failure or Cirrhosis [20]. A total of 885 patients were enrolled into the programme in the study time, from which 591 were included in this study, and 294 were censored.

Variables

The dependent variable for this study is defined as the occurrence of death within the first year that a patient spends in the high-risk programme; in other words, 1-year mortality since enrolment.

A broad range of predictor variables were incorporated into the analysis. These are not disease-specific variables, but general criteria that can be applied to a heterogeneous population of chronic patients with different combinations of illnesses.

These include demographic data (age, sex, patient living in nursing home or not, presence of informal carer), functional status variables (Barthel scale [21], Lawton & Brody scale [22], and degree of dependency), nutritional values (serum albumin), renal deterioration indicators (creatinine, albumin/creatinine index), prevalence of selected non-cancer conditions, total number of selected comorbidities, prevalence of pressure skin ulcers, intake of opioids and psycholectics, total number of active prescriptions and GMA score.

All variables represent real world data and were collected at patient inclusion into the high risk programme. At that time, all patients undergo an assessment in which doctors, nurses and social workers classify patients into four levels of severity (low, moderate, high and very high). All three classifications are then combined in a meeting in which all healthcare professionals involved in the assessment produce a 'global status' variable, with the same levels of severity [8].

Multiple Imputation was used to fill in missing values, under a missing at random (MAR) assumption, with the purpose of avoiding any possible bias that incomplete data would have caused [23].

Statistical Analysis

Descriptive statistics are provided including the mean and standard deviation for continuous variables, and number of people and percentage for categorical variables, both for the whole population, and by 1-year mortality status.

A logistic regression model was fitted following the methodology defined by Hosmer & Lemeshow [24] and Steyerberg [25]. Univariate analysis was performed on each variable as a first step of model selection, showing the level of association between 1-year mortality and every predictor variable. When a variable had a p-value significance level of $p \leq 0.25$, it was considered further for multivariate modelling. Subsequently, a model containing all of the pre-selected variables from the univariate analysis was fit, and insignificant variables were deleted one by one in separate steps until a final model was reached.

Discriminative ability of the model was measured using the Receiver Operating Characteristic (ROC) curve and its associated Area Under the Curve (AUC), which evaluates discrimination of the model across all probability cut-offs that classify the output as positive or negative. Whenever two variables measured the same aspect (e.g.

Barthel scale vs degree of dependency) the variable that yielded a higher AUC was selected for the final model.

Regarding goodness of fit, the Hosmer-Lemeshow test and calibration plots were used. To validate the model, we used k-fold Cross-Validation with 10 random Bootstrap Validation, rather than the traditional method of splitting the dataset into developing and validation samples, which has proved to be inefficient in comparison to the methods mentioned above [25]. Results are expressed in terms of odds ratios (OR), and analyses were performed using STATA 15.0.

To favour the applicability of the prediction model and the introduction of individual risk assessment in clinical practices, we provide a nomogram that was created with the 'nomolog' command in STATA [26]. This graphical calculation instrument uses the results of our model and constitutes an easy and rapid way of providing individual estimations of the probability of death for each patient.

RESULTS

Table 1 shows descriptive statistics of all predictor variables considered in analysis, for the whole sample, and by 1-year mortality status. This table also shows univariate logistic regression p-values, showing level of significance of the relationship between each variable with 1-year mortality. The profiles of those who die within one year after inclusion of the program and those who live can be extracted from this table.

There is a significant difference in age between those who died and those who survived the follow-up, but sex is not a significant variable, with similar proportions of males and females in both groups. Living in a nursing home, or having an informal carer did not have any effects on death.

Functional status variables were good prognostic indicators of death, with the Barthel Scale and the Degree of Dependency showing a strong association with mortality. However when the Lawton & Brody Scale was used, functional deterioration was not a significant variable.

Table 1 Descriptive Statistics and Univariate Logistic Regression Results

Variable	All patients n=591	1-Year Mortality		Univariate Logistic Odds Ratios	Univariate Logistic p-values
		Alive n=390 (66%)	Dead n=201(34%)		
Age	83.60 ±8.10	82.74 ±8.47	85.25 ±7.07	1.043	<0.001
Sex					
Female	255 (43.15%)	170 (43.59%)	85 (42.29%)		
Male	336 (56.85%)	220 (56.41%)	116 (57.71%)	1.054	0.762
Living in Nursing Home	12 (2.03%)	9 (2.31%)	3 (1.49%)	0.641	0.509
Presence of Informal Carer	532 (90.02%)	346 (88.72%)	186 (92.54%)	1.539	0.295
Barthel Scale	58.13 ±30.2	65.04 ±28.42	44.74 ±29.11	0.978	<0.001
Lawton & Brody Scale	3.16 ±1.81	3.29 ±1.86	2.89 ±1.68	0.913	0.125
Degree of Dependency					
Not dependent	112 (18.95%)	96 (24.62%)	16 (7.96%)		
Degree I: Moderate	295 (49.92%)	204 (52.31%)	91 (45.27%)	1.780	0.058
Degree II: Severe	123 (20.81%)	60 (15.38%)	63 (31.34%)	3.705	<0.001
Degree III: Great Dependency	61 (10.32%)	30 (7.69%)	31 (15.42%)	4.537	<0.001
Serum Albumin (g/dL)	3.79 ±0.43	3.85 ±0.41	3.67 ±0.43	0.422	<0.001
Creatinine (mg/dL)	1.55 ±0.81	1.51 ±0.76	1.63 ±0.89	1.193	0.094
Albumin/Creatinine Index					
≤30 mg/g: Normal	359 (60.74%)	238 (61.02%)	121 (60.20%)		
30-300 mg/g: Moderate	167 (28.26%)	112 (28.72%)	55 (27.36%)	0.913	0.653
≥300mg/g: High	65 (11.00%)	40 (10.26%)	25 (12.44%)	1.256	0.419
Prevalence of Diabetes	428 (72.42%)	284 (72.82%)	144 (71.64%)	0.943	0.761
Prevalence of Chronic Renal Failure	410 (69.37%)	274 (70.26%)	136 (67.66%)	0.886	0.517
Prevalence of Ischemic Heart Disease	293 (49.58%)	195 (50.00%)	98 (48.76%)	0.951	0.775
Prevalence of Heart Failure	402 (68.02%)	257 (65.90%)	145 (72.14%)	1.340	0.124
Prevalence of Cerebrovascular Disease	198 (33.50%)	127 (32.56%)	71 (35.32%)	1.131	0.501
Prevalence of COPD	176 (29.78%)	122 (31.28%)	54 (26.87%)	0.807	0.266
Prevalence of Asthma	123 (20.81%)	84 (21.54%)	39 (19.40%)	0.877	0.545
Prevalence of Dementia	110 (18.61%)	65 (16.67%)	45 (22.39%)	1.442	0.091
Prevalence of Cirrhosis	46 (7.78%)	33 (8.46%)	13 (6.47%)	0.748	0.393
Number of Selected Comorbidities	3.70 ±0.84	3.69 ±0.85	3.71 ±0.82	1.017	0.873
Presence of Pressure Ulcers	187 (31.64%)	103 (26.41%)	84 (41.79%)	2.000	<0.001
Intake of Opioids	77 (13.03%)	54 (13.85%)	23 (11.44%)	0.804	0.412
Intake of Psycholectics	54 (9.14%)	31 (7.95%)	23 (11.44%)	1.496	0.165
Number of Active Prescriptions	8.05 ±3.51	8.07 ±3.53	8.03 ±3.48	0.997	0.904
GMA Score	23.50 ±6.72	23.69 ±6.68	23.14 ±6.79	0.987	0.342
Clinical Severity					
Mild	2 (0.34%)	2 (0.51%)			
Moderate	69 (11.68%)	54 (13.85%)	15 (7.46%)		
Severe	257 (43.49%)	178 (45.64%)	79 (39.30%)	0.901	0.578
Very Severe	263 (44.50%)	156 (40.00%)	107 (53.23%)	7.902	<0.001
Nursing Needs					
Mild	47 (7.95%)	40 (10.26%)	7 (3.48%)		
Moderate	204 (34.52%)	160 (41.03%)	44 (21.89%)	1.493	0.383
Severe	213 (36.04%)	125 (32.05%)	88 (43.78%)	3.503	0.005
Very Severe	127 (21.49%)	65 (16.67%)	62 (30.85%)	5.120	<0.001
Social Needs					
Mild	233 (39.42%)	152 (38.97%)	81 (40.30%)		
Moderate	280 (47.38%)	181 (46.41%)	99 (49.25%)	0.931	0.726
Severe	67 (11.34%)	51 (13.08%)	16 (7.96%)	0.616	0.189
Very Severe	11 (1.86%)	6 (1.54%)	5 (2.49%)	1.502	0.489
Global Status					
Mild	2 (0.34%)	2 (0.51%)			
Moderate	86 (14.55%)	76 (19.49%)	10 (4.98%)		
Severe	474 (80.20%)	306 (78.46%)	168 (83.58%)	4.220	<0.001
Very Severe	29 (4.91%)	6 (1.54%)	23 (11.44%)	29.082	<0.001

Dependent Variable: 1-Year Mortality; COPD: Chronic Obstructive Pulmonary Disease; Scale Ranges: Barthel: 0–100, Lawton & Brody:0-8

Individuals that died showed lower nutritional values as measured by serum albumin, and higher creatinine values, indicating renal deterioration. In contrast, no significant differences were reported in the albumin/creatinine index. Pressure ulcers were much more frequent in those that died, with significant differences.

Regarding prevalence of non-cancer indications, those patients with dementia showed higher association with mortality. All other indications showed similar prevalence across groups, and no significant differences were observed in number of selected comorbidities.

Both groups had a similar number of active prescriptions at the time of inclusion into the programme, and no differences in the intake of opioids and psycholectics were observed. The GMA score was not a significant variable when predicting death.

Significant differences between groups were observed in the professional-rated variables, such as clinical severity (only at the very severe classification level), nursing needs (at the severe and very severe classification levels) and global status. However, social needs were similar for all patients.

Results from descriptive statistics and univariate logistic analyses can inform about the representative profile of those high-risk patients that die within 1 year follow-up. In general, they are elderly patients (≈85 years old) being cared by a 3rd person, with functional, nutritional and renal decline. They are clinically severe individuals with high nursing care needs, but with the same social needs as those that do not die. The most widespread conditions among this type of patients are heart failure (72.14%), diabetes (71.64%) and chronic renal failure (67.66%). On the other hand, asthma (19.4) and cirrhosis (6.47%) are less common.

With respect to multivariate modelling (Table 2), variables found to be independently associated to death and included into the final prediction model (Model 1) were the Barthel Scale, creatinine value, and patient global status. When considering the Degree of Dependency as a measure of functional status (Model 2) rather than the Barthel Scale, the existence of pressure ulcers was also fit into the final model. However the model that included the Barthel Scale was considered the most appropriate model. Some additional basic models are included in a supplementary appendix.

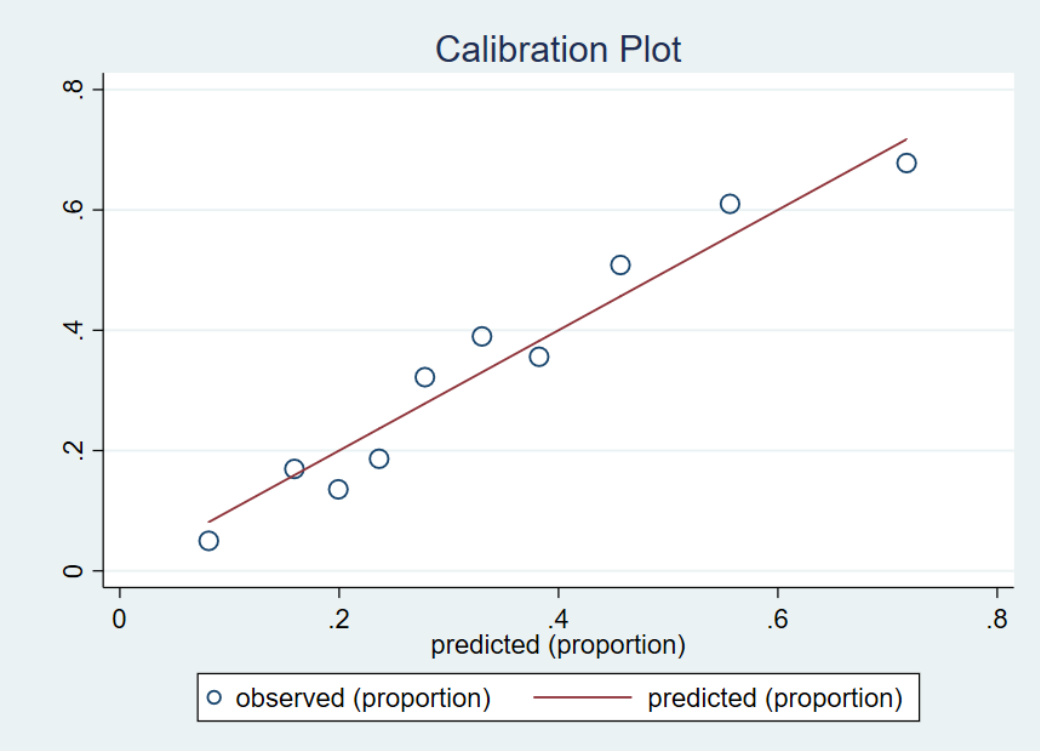
Table 2 Multivariate Logistic Regression to 1-Year Mortality

	Model 1		Model 2	
Variable	**Odds Ratio (95% CI)**	**p-value**	**Odds Ratio (95% CI)**	**p-value**
Barthel Scale	0.979 (0.972 - 0.987)	0.000	-	-
Degree of Dependency II: Severe	-	-	2.343 (1.203 - 4.565)	0.016
Degree of Dependency III: Great Dependency	-	-	2.642 (1.465 - 4.764)	0.001
Creatinine (mg/dL)	1.284 (1.022 - 1.613)	0.032	1.290 (1.030 - 1.616)	0.027
Presence of Pressure Ulcers	-	-	1.572 (1.048 - 2.358)	0.029
Severe Global Status	3.573 (1.752 - 7.286)	0.000	3.829 (1.874 - 7.823)	0.000
Very Severe Global Status	20.699 (6.408 - 66.855)	0.000	21.846 (6.797 - 70.215)	0.000
Constant	0.322 (0.140 - 0.737)	0.007	0.057 (0.026 - 0.126)	0.000
Area Under the Curve	0.751		0.737	

Dependent Variable: 1-Year Mortality; CI: Confidence Interval

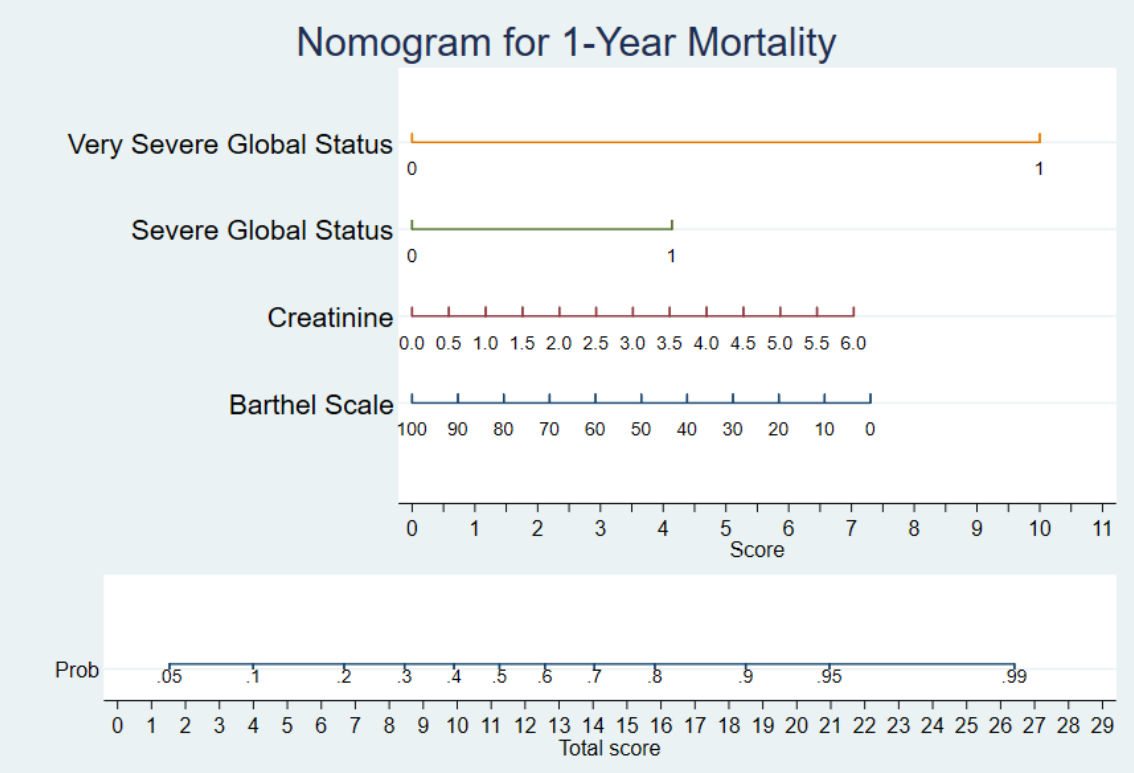
The Hosmer-Lemeshow calibration test returned a p-value of 0.761, and the calibration plot is presented in Figure 1. The model discrimination capability, measured as the area under the ROC curve, is 0.751. Regarding validation, the Cross-Validation method yielded an AUC of 0.744 (0.701, 0.788 C.I.), and when Bootstrapping was used, the AUC was 0.751 (0.711, 0.791 C.I.). The percentage of correctly classified patients is 72.08%, with a sensitivity of 57.21% and a specificity of 79.74%.

Figure 1: Model calibration plot



A graphical calculation instrument, a nomogram, is provided in Figure 2. This tool synthesizes the results from the final multivariate model and provides a way to estimate individual probabilities of dying. The nomogram uses the observed values for each variable of interest and produces an estimated probability of death for each patient through a set of interconnected scales.

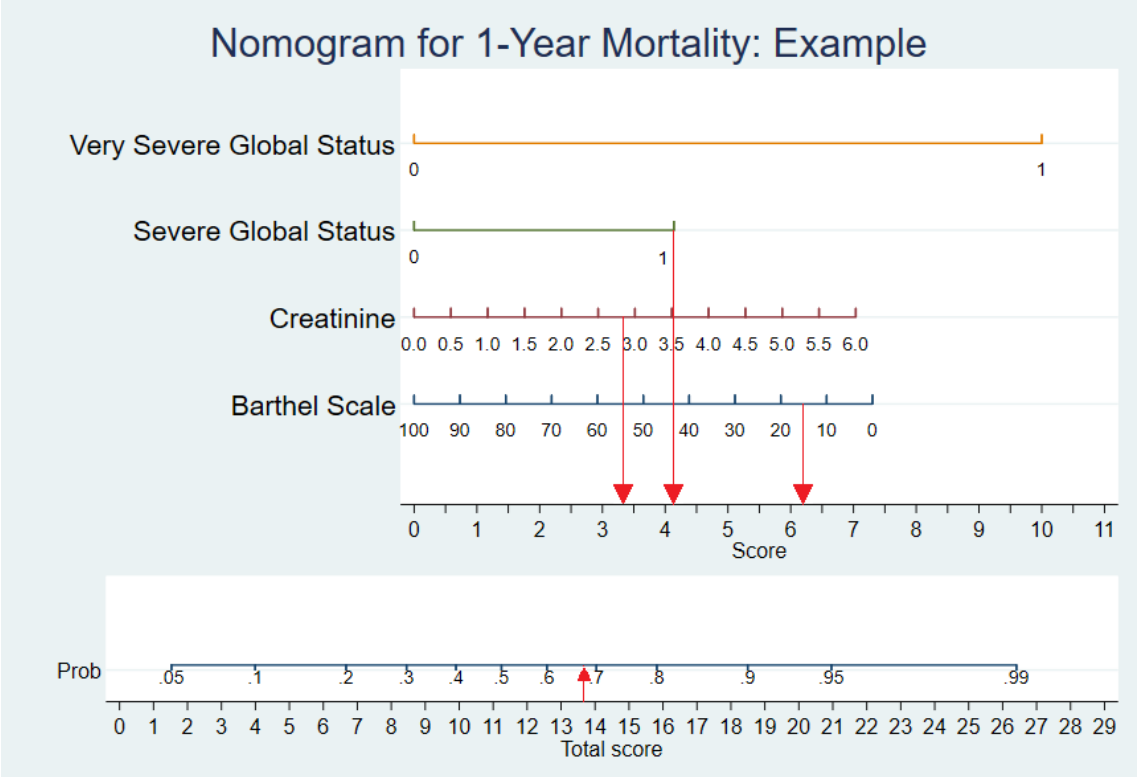
Figure 2: Nomogram for graphical risk score calculation



Consider the following example (Figure 3) on how to use the provided nomogram. We estimate the probability of death for a patient with the following observed values for identified predictors: Barthel Scale: 15; Creatinine: 2.86 mg/dL; Global Status: Severe. These values correspond to the following risk scores: Barthel Scale ≈ 6.2 points; Creatinine ≈ 3.3 points; Severe Global Status ≈ 4.2. The total risk score for this patient is 6.2+3.3+4.2 = 13.7, which corresponds approximately to a probability of dying during 1-year follow-up ≈ 67%.

The nomogram can be used to apply the model estimated above to all patients at patient enrolment, and at any time thereafter if patient conditions change or regular re-assessments are scheduled. New data collection for variables included into the final prediction model would be required when re-assessments are performed.

Figure 3: Example of nomogram risk score calculation



DISCUSSION

This study developed and validated a mortality prediction model that estimates risk of death and identifies independent predictors of death for a population of non-cancer multiple chronic conditions patients with a level of discrimination power equivalent to other mortality prediction models with similar sample sizes and mortality rates.

Several patient features were recognised as relevant prognostic indicators of death. These include functional status, renal deterioration, pressure ulcers and patient global status, which measures three different dimensions: clinical severity, nursing needs and social needs.

Concerning functional status, the Barthel Scale was a better predictor in comparison to the Degree of Dependency in AUC terms. Besides, whereas the Barthel score can be easily calculated, the Degree of Dependency requires an award by a qualified committee [16], making the proceeding slow and time consuming. Therefore using the model that includes the Barthel Scale is not only better when considering AUC, but also simpler, faster, and easier to use in clinical settings.

Global status was also preferred to the individual inclusion of clinical severity, nursing needs or social needs variables on the basis of higher AUC values, but also because global status is a combination of all the former, and therefore taking all three dimensions into account at the same time.

The recognition of these predictors constitutes the starting process of early identification of palliative patients' needs [15]. Our model helps in detecting those needs and guiding transitions towards end-of life care. However, mortality prediction presents a still image of patient prospects. Frequent re-assessments should monitor changing progression of each patient's needs. Patient decline would be observed, and care strategies would be tailored to changing needs.

Among study limitations, cognitive status variables were not included in the analysis. In our sample, cognitive status variables such as the Mini-Mental State Examination are being collected predominantly in those patients diagnosed with dementia, resulting in very high missing data rates (73%) that would make results unreliable and biased. Therefore, although these variables were collected, we decided not to include them in the analysis. Furthermore, we lacked quality of life data, for instance the EQ-5D, which might have improved risk prediction and helped to achieve higher AUC values. To our knowledge, no other mortality prediction models that focus on chronic patients have included such variables.

Some issues can affect transferability of the study. Some predictor variables may not be available in certain contexts, as it is the case for professional-rated variables, that may be registered in different ways. Degree of dependency is a variable that is registered according to Spanish law, representing an additional reason to use the model that includes the Barthel score as a predictor. Even so, using degree of dependency is still worth including as it shows the effect of pressure ulcers in model 2. GMA score is also widely available in Spain but not in other countries, although the top 5% risks of the population can still be identified using other stratification tools. As a final point, the model was not validated using an external dataset, but using internal validation techniques instead.

In relation to the discrimination ability of our model, a systematic review by Siontis et al. [27] showed that mortality prediction models with sample sizes between 288 to 810 participants had an AUC of 0.76 on average, and if mortality rates were above 33%,

mean AUC values were 0.73. Given these data we considered that our model had a good degree of discrimination ability; especially taking into account the specific characteristics of multiple chronic conditions patients such as entry-re-entry life trajectories, which make death difficult to predict.

Similar mortality prediction models for chronic patients have been carried out in the past, however these included cancer patients [28, 29], or they included patients with no comorbidities [29, 30, 31].

Despite these differences in study populations, functional status and the Barthel Scale in particular was identified as a key indicator of death across the literature, confirming the results presented above [28,31]. Badia et al. [31] also confirmed that pressure ulcers have a significant role in end of life. These studies also found that nutritional values [15] and dementia [28] were strong predictors of death. However we only found serum albumin and dementia significant at the univariate analysis.

Our results showed that age is related to death, but not as an independent predictor. The reason for this result is that age and functional status were closely correlated between them, and thus only functional status was included into the final model. While some authors confirmed our findings [29], others included age in their models [28].

Concerning GMA score, Dueñas-Espin et al. [18] found that it is a good predictor in explaining healthcare expenditure and resource utilisation, yet it does not perform as well when explaining mortality. These judgments are supported by our results.

No standard definition of multiple chronic conditions patients is given in the literature, with some authors taking into consideration patients with 2 or more chronic conditions [3, 28], and others [32] (including us) defining them as patients with 3 or more simultaneous conditions. Moreover, the selection of clinical conditions included in analysis varies from study to study. The programme considered in this study only allows enrolment of patients with three or more pre-defined chronic illnesses, according to a previous segmentation based on number of comorbidities, severity and age [20]. Patients that do not meet these criteria are not enrolled in the programme, leaving a considerable amount of patients with less or other non-cancer chronic conditions out of it. Since these patients fall out of the scope of action of the programme, we do not have data to estimate how well the model predicts mortality on them. This remains as a future area of research.

CONCLUSION

With the purpose of providing tools for the identification of patients that would benefit from non-cancer palliative care rather than integrated care programmes, a mortality predictive model has been developed in a population of multiple chronic conditions patients. Results have been translated into a nomogram to enhance their applicability in clinical practices. In this way, probability of death can be estimated for each individual to allow flagging of potential palliative care patients.

REFERENCES

1. Moffat K, Mercer SW, Challenges of managing people with multimorbidity in today's healthcare systems, BMC Family Practice, 2015, 16:129, DOI 10.1186/s12875-015-0344-4
2. Ward BW, Schiller JS, Goodman RA. Multiple Chronic Conditions Among US Adults: A 2012 Update. Prev Chronic Dis 2014;11:130389. DOI:doi.org/10.5888/pcd11.130389
3. Lenzi J, Avaldi VM, Rucci P et al. Burden of multimorbidity in relation to age, gender and immigrant status: a cross-sectional study based on administrative data. BMJ Open 2016; 6: e012812. doi: 10.1136/bmjopen-2016-012812
4. Hajat C, Stein E, The global burden of multiple chronic conditions: A narrative review, Preventive Medicine Reports, 2018, Vol.12, 284-293, doi.org/10.1016/j.pmedr.2018.10.008.
5. Orueta JF, Nuño-Solinis R, Mateos M et al. Predictive risk modelling in the Spanish population: a cross-sectional study. BMC Health Services Research, 2013, 13:269. doi: 10.1186/1472-6963-13-269
6. Sambamoorthi U, Tan X, Deb A. Multiple chronic conditions and healthcare costs among adults Expert Rev Pharmacoecon Outcomes Res. 2015;15(5):823–832. doi:10.1586/14737167.2015.1091730
7. Lanzeta I, Mar J, Arrospide A. Cost-utility analysis of an integrated care model for multimorbid patients based on a clinical trial; Gac Sanit, 2016; 30(5): 352-358. doi.org/10.1016/j.gaceta.2016.05.002.
8. Ibarrola Guillén C, Librero López J, Celorrio Astiz J.J et al. Evaluación Estrategia Navarra de Atención Integrada a pacientes pluripatológicos a los dos años de implantación; 2018, Depósito legal DL NA 668/2018. Available at www.navarra.es Accessed September 2019
9. Coventry PA., Grande GE., Richards DA et al. Prediction of appropriate timing of palliative care for older adults with non-malignant life-threatening disease: a systematic review, Age and Ageing, 2005, Volume 34, Issue 3, 218–227 doi.org/10.1093/ageing/afi054
10. Alonso-Babarro A. Aguilar, A Rexach, L. Selection criteria of patients with non-oncological diseases in palliative care programs and/or services, 2010, Medicina Paliativa. 17. 161-171.
11. GBD 2015 Mortality and Causes of Death Collaborators, Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980-2015: a systematic analysis for the Global Burden of Disease Study, Lancet, 2016, vol. 388,10053: 1459-1544 doi:10.1016/S0140-6736(16)31012-1
12. Hess S, Stiel S, Hofmannet S et al. Trends in specialized palliative care for non-cancer patients in Germany - Data from the National Hospice and Palliative Care Evaluation (HOPE) European Journal of Internal Medicine, 2014, Volume 25, Issue 2, 187 – 192 doi.org/10.1016/j.ejim.2013.11.012
13. Grbich C, Maddocks I, Parker D et al. Identification of patients with noncancer diseases for palliative care services. Palliative and Supportive Care, Cambridge University Press; 2005; 3(1):5–14. doi.org/10.1017/S1478951505050029
14. Thai V, Ghosh S, Tarumi Y, et al. Clinician prediction survival of end stage non-cancer patients, Progress in Palliative Care, 2016, doi: 10.1080/09699260.2016.1159416

15. Amblàs-Novellas J, Murray SA, Espaulella J, et al. Identifying patients with advanced chronic conditions for a progressive palliative care approach: a cross sectional study of prognostic indicators related to end-of life trajectories, BMJ Open, 2016; 6:e012340. doi: 10.1136/bmjopen-2016-012340
16. Ley 39/2006, de 14 de diciembre, de Promoción de la Autonomía Personal y Atención a las personas en situación de dependencia. Boletín Oficial del Estado, de 15/12/2006. Available at www.boe.es Accessed September 2019
17. Monterde, D., Vela, E., Clèries, M et al. Los grupos de morbilidad ajustados: nuevo agrupador de morbilidad poblacional de utilidad en el ámbito de la atención primaria. Atención Primaria, 2016, 48(10), 674-682. doi: 10.1016/j.aprim.2016.06.003
18. Dueñas-Espín I, Vela E, Pauws S, et al. Proposals for enhanced health risk assessment and stratification in an integrated care scenario, BMJ Open, 2016; 6:e010301. doi: 10.1136/bmjopen-2015-010301
19. Spanish Ministry of Health, Consumption and Social Welfare (Ministerio De Sanidad, Consumo Y Bienestar Social) Informe del proyecto de Estratificación de la Población por Grupos de Morbilidad Ajustados (GMA) en el Sistema Nacional de Salud (2014-2016); 2018. Available at: https://www.mscbs.gob.es/organizacion/sns/planCalidadSNS/pdf/informeEstratificacionGMASNS_2014-2016.pdf ; Accessed March 2020
20. Ibarrola Guillén C., Sanz Barea Ángel., Iribarren Udobro I., Estrategia Navarra De Atención Integrada a Pacientes Crónicos y Pluripatológicos, 2016, Available at www.navarra.es Accessed March 2020
21. Mahoney FI, Barthel DW. Functional evaluation: the Barthel Index. Md State Med J, 1965; 4:61–5.
22. Lawton MP, Brody EM. Assessment of older people: self maintaining and instrumental activities of daily living. Gerontologist 1969;9:179–86
23. White IR., Royston P, Wood AM, Multiple imputation using chained equations: Issues and guidance for practice. Statist; Med. 2011, 30: 377-399. doi: 10.1002/sim.4067
24. Hosmer DW, Lemeshow S, Applied Logistic Regression, Second Edition, John Wiley & Sons, 2000, ISBN: 9780471356325; DOI: 10.1002/0471722146
25. Steyerberg EW. Clinical Prediction Models: A Practical Approach to Model Development, Validation, and Updating, Springer, 2009, ISBN: 978-0-387-77243-1; DOI: 10.1007/978-0-387-77244-8
26. Zlotnik, A. Abraira, V. A General-purpose Nomogram Generator for Predictive Logistic Regression Models. The Stata Journal, 2015, 15(2), 537–546. https://doi.org/10.1177/1536867X1501500212
27. Siontis GCM, Tzoulaki I, Ioannidis JPA, Predicting Death: An Empirical Evaluation of Predictive Tools for Mortality. *Arch Intern Med.* 2011; 171(19):1721–1726. doi:10.1001/archinternmed.2011.334
28. Bernabeu-Wittel M, Ollero-Baturone M, Moreno-Gaviño L et al, Development of a new predictive model for polypathological patients. The PROFUND index, Eur J Intern Med, 2010, doi:10.1016/j.ejim.2010.11.012
29. Louis DZ, Robeson M, McAna J, et al. Predicting risk of hospitalisation or death: a retrospective population-based analysis, BMJ Open, 2014; 4: e005223. doi: 10.1136/bmjopen-2014-005223
30. Orueta JF, García Álvarez A, Aurrekoetxea JJ, et al. FINGER (Forming and Identifying New Groups of Expected Risks): developing and validating a new

predictive model to identify patients with high healthcare cost and at risk of admission, BMJ Open, 2018; 8:e019830. doi: 10.1136/bmjopen-2017-019830
31. Badia JG, Borràs Santos A, Contel Segura JC et al, Predictors of mortality among elderly dependent home care patients. BMC Health Services Research, 2013, 13:316. doi: 10.1186/1472-6963-13-316
32. Aira Ferrer P, Domínguez Lorenzo C, Cano González JM et al. Características de los pacientes con enfermedad crónica avanzada incluidos en un programa de cuidados paliativos domiciliario; Med Paliat. 2015. doi.org/10.1016/j.medipa.2015.08.001

CHAPTER 2

RISK STRATIFYING THE MULTIPLE CHRONIC CONDITIONS PATIENTS POPULATION: A CLUSTER ANALYSIS APPROACH

Published as:

"Tailoring Integrated Care Services for High-Risk Patients with Multiple Chronic Conditions: A Risk Stratification Approach Using Cluster Analysis"

at BMC Health Services Research

Citation:

Bretos-Azcona PE, Sánchez-Iriso E, Cabasés Hita JM. Tailoring integrated care services for high-risk patients with multiple chronic conditions: a risk stratification approach using cluster analysis. BMC Health Serv Res. 2020 Aug 27;20(1):806. doi: 10.1186/s12913-020-05668-7.

Presented at the conferences:

XXXIX Jornadas de la Asociación de Economía de la Salud (12-14 June 2019); Albacete, Spain

ABSTRACT

Background

The purpose of this study was to produce a risk stratification within a population of high-risk patients with multiple chronic conditions who are currently treated under a case management program and to explore the existence of different risk subgroups. Different care strategies were then suggested for healthcare reform according to the characteristics of each subgroup.

Methods

All high-risk multimorbid patients from a case management program in the Navarra region of Spain were included in the study (n=885). A 1-year mortality risk score was estimated for each patient by logistic regression. The population was then divided into subgroups according to the patients' estimated risk scores. We used cluster analysis to produce the stratification with Ward's linkage hierarchical algorithm. The characteristics of the resulting subgroups were analyzed, and post hoc pairwise tests were performed.

Results

Three distinct risk strata were found, containing 45%, 38% and 17% of patients. Age increased from cluster to cluster, and functional status, clinical severity, nursing needs and nutritional values deteriorated. Patients in cluster 1 had lower renal deterioration values, and patients in cluster 3 had higher rates of pressure skin ulcers, higher rates of cerebrovascular disease and dementia, and lower prevalence rates of chronic obstructive pulmonary disease.

Conclusions

This study demonstrates the existence of distinct subgroups within a population of high-risk patients with multiple chronic conditions. Current case management integrated care programs use a uniform treatment strategy for patients who have diverse needs. Alternative treatment strategies should be considered to fit the needs of each patient subgroup.

Keywords: Risk stratification, Integrated care, Case management, Cluster Analysis

BACKGROUND

Decades of progressive declines in the burden of communicable diseases and consequent improvements in life expectancy have shifted clinical and managerial concerns towards chronic illnesses, which are reaching alarming levels of prevalence in aging societies [1,2].

Special attention has been given to multimorbidity [3,4] and high-risk multiple chronic condition (MCC) patients in particular. Despite representing a small share of the chronic patient population, high-risk MCC patients account for a great share of healthcare organization budgets [5]. The elevated number of consultations, hospitalizations, and other treatments from different, uncoordinated specialties decreases favorable outcomes and increases cost [6].

Plans for appropriate management and delivery of care in the context of high-risk patients focus efforts on the realignment of systems towards case management integrated care programs [1,7]. These models plan and coordinate care around specific high-risk patients through the assignment of a reference physician or a small multidisciplinary team. Teams assess the individual needs of each patient, develop a care plan accordingly and coordinate treatment delivery. Patients are monitored with periodic reassessments [8].

Identifying patients for which case management would be appropriate is an essential element of programs of this nature, and it is usually done by means of risk stratification techniques that classify patients with similar clinical needs into homogeneous groups [9]. This requires the establishment of a risk score using statistical models together with judgments from clinicians and the formation of certain thresholds for the assignment of patients to different risk strata [10].

In general terms, candidates for case management belong to the top 5% risk stratum of the population and are identified using a variety of ready-to-use risk stratification tools, including clinical risk groups (CRGs), adjusted clinical groups, diagnosis-related groups, diagnostic cost groups or the senior segmentation algorithm among others [9, 11-13]. A set of common services can be provided where risk stratification produces a homogeneous group of patients, and if needs are appropriately addressed, case management will fit patients in a cost-effective way, avoiding wasteful, unnecessary care.

However, evidence has shown that case management programs are not cost-effective in comparison to non-integrated care programs for high-risk patients [7,8,14]. Case management interventions are not suitable for all high-risk patients but for a subset of patients who would benefit from them [14]. In other words, the population they target is heterogeneous and has different needs [10,15-17], yet all patients are treated in a uniform manner under the same case management strategy. Since some groups of patients are receiving a type of care that does not fit their needs, care provides minimal or no health benefit to those patient subgroups and does not justify the costs, translating into low-value care for some of the patients in the high-risk population [18].

Therefore, case management requires further stratification to identify those patient subgroups that do not benefit from their current care and to adapt care strategies for them. The purpose is to target appropriate care for the appropriate patients. By reorganizing high-risk integrated care programs, we aim to target new services to selected groups of patients that are most likely to benefit from them. The extent to which newly organized services fit the clinical needs of patient subgroups will determine both improvements in outcomes and the degree of efficiency in healthcare resource utilization.

The purpose of this study was to produce a restratification within the high-risk MCC patient population, exploring the existence of different risk subgroups. Subsequently, the characteristics of each risk stratum were defined. Finally, we proposed different care strategies according to the risk profile of each subgroup, tailoring integrated care for high-risk patients.

METHODS

Data and participants

In 2016, the Navarra region of Spain implemented an integrated care program for the treatment of chronic illness, which included a case management model for high-risk, noncancer MCC patients [19]. This study included all high-risk MCC patients who were treated in the region's case management program from April 2016 – August 2018. The conditions to qualify for program enrollment were as follows:

Patients suffered at least three selected noncancer pathologies, including heart failure, dementia, ischemic heart disease, cerebrovascular disease, diabetes, chronic obstructive

pulmonary disease, asthma, chronic renal failure and cirrhosis, and patients belonged to the top 5% of the risk pyramid according to the adjusted morbidity groups (GMA). GMA is a stratification tool similar to CRGs that is widely applied in Spain [20,21]. A total of 885 patients were considered.

Data were obtained from the high-risk case management program database, which is anonymized and includes sociodemographic data, as well as data regarding functional status (Barthel score), nutritional status (serum albumin), renal deterioration status (creatinine, albumin/creatinine index), the presence of pressure skin ulcers, the number of prescriptions, prevalence and number of coexisting selected illnesses and the GMA risk score. In addition, the database also incorporates professionally rated variables such as clinical severity, nursing needs and social needs. A combination of the former is also available as global severity status. All variables were measured at patient inclusion in the program, when patients underwent a comprehensive assessment of their situation. Missing values were filled using multiple imputation to avoid biases in risk score estimation and subsequent stratification (Table 2) [22].

Producing a risk score

A risk score was estimated for each of the patients using data from the initial comprehensive assessment that was completed upon inclusion in the case management program. The outcome for this risk score estimation was 1-year mortality from enrollment in the case management program. The reason why mortality was used as the outcome was that our population of interest consisted of patients with different chronic illness combinations. Therefore, disease-specific outcomes were not appropriate, as it was not possible to apply them to all patients under study. A common outcome was needed, and 1-year mortality was selected.

The risk score was estimated by logistic regression, where we first tested all variables in univariate analyses. Those variables that were significant were then fitted into a multivariate model, and insignificant variables were eliminated from the model in a stepwise manner. Significant predictors included the functional status, creatinine value, global severity status and presence of pressure skin ulcers. The results were validated using cross-validation techniques, as well as bootstrapping. A full description of the risk score estimation process is available in another published study [23]. Subsequently, patients were categorized into different ‘buckets’ or clusters according to their estimated risk [24].

Risk stratification

For the purpose of determining patient subgroups and categorizing individuals into distinct, mutually exclusive risk strata, we used machine learning algorithms. These techniques group individuals who have similar risk scores into subgroups that are dissimilar and are more frequently termed cluster analysis [25].

We used Ward's linkage hierarchical algorithm with the squared Euclidian distance (L2squared). The optimal number of clusters was determined using the Duda/Hart stopping rule and visually through a dendrogram (Figure 1). A dendrogram is a diagram that shows how observations merge with other observations that are similar to them in terms of distance. Those observations that are closest to them, or equivalently those that have the most similar risk scores, are merged into a group. This process continues iteratively, and larger, distinct groups can be observed in the dendrogram. Mergers are represented as nodes, and the distance between groups of patients is shown in the vertical axis. The results were validated using silhouettes, reassigning individuals to a different cluster when needed [26].

To evaluate the stability of the results, the full sample was randomly divided into four equally sized subsamples, each containing 25% of the observations, and the algorithm was run again on each subsample. It is possible to think about this process as k-fold cross validation with four folds. The robustness of the results was further tested by running the K-means algorithm, setting the parameter k equal to Ward's linkage optimal number of clusters. Table 1 shows how many patients belong to each cluster when using the two clustering techniques used in this study, Ward's linkage algorithm, and K-means algorithm that was performed as a robustness check. In addition, the number of patients in each of the four randomly divided subsamples from stability analysis is shown.

Cluster examination

Following the identification of patient subgroups, their clinical and sociodemographic characteristics were compared to test if there were significant differences between them. The prevalence of chronic illness diagnoses and their most frequent combinations were also compared.

When considering continuous variables, one-way ANOVA tests were performed. Kruskal-Wallis tests were used when considering categorical variables or if the assumptions for ANOVA did not hold, and a $\chi 2$ test was used for binary variables. If

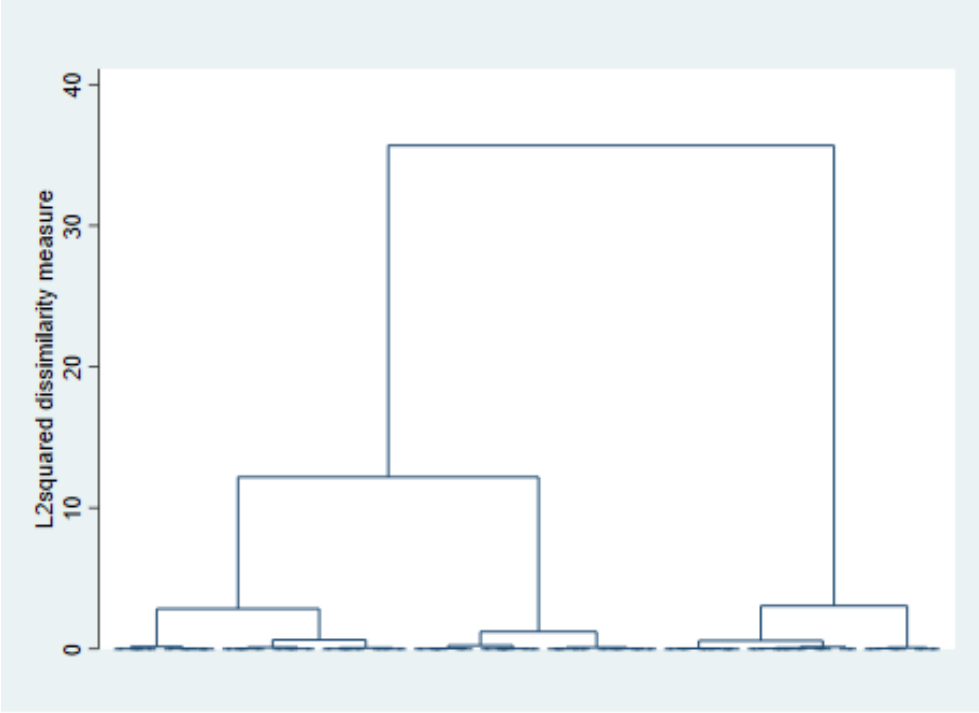
significant differences in patient characteristics were observed across clusters, further post hoc pairwise tests were completed to detect which cluster was different from the remaining clusters. Multiple one-way ANOVA comparisons with Bonferroni corrections, Mann-Whitney U tests and Fisher's exact tests were used if ANOVA, Kruskal-Wallis or $\chi 2$ tests were used, respectively. All analyses were carried out using STATA 15.0 software.

RESULTS

Clustering results

Both the Duda/Hart stopping rule and the clustering process dendrogram (Figure 1), which shows the last 100 grouping nodes, indicated the presence of three distinct clusters within the high-risk MCC patient population. The optimal number of clusters was also assessed considering its clinical relevance and interpretability.

Figure 1: Patient grouping process (dendrogram)



The clusters were ordered in such a way that the 1-year mortality risk scores were incremental. Hence, patients with the lowest risk scores belong to cluster 1, cluster 2 includes intermediate cases, and very high-risk patients have been allocated to cluster 3. With respect to the distribution of patients among clusters, cluster 1 included ≃45% of patients, ≃38% were classified into cluster 2, and ≃17% were assigned to cluster 3.

Table 1: Distribution of patients among clusters

	Cluster 1	Cluster 2	Cluster 3
Ward's linkage (n=885)	400 (45.20%)	336 (37.96%)	149 (16.84%)
Stability Analysis			
Sub-Sample 1 (n=221; 25%)	91(41.18%)	91(41.18%)	39(17.64%)
Sub-Sample 2 (n=221; 25%)	98 (44.34%)	89 (40.28%)	34 (15.38%)
Sub-Sample 3 (n=221; 25%)	105 (47.52%)	74 (33.48%)	42 (19.00%)
Sub-Sample 4 (n=222; 25%)	100 (45.05%)	88 (39.64%)	34 (15.32%)
Total (n=885)	394 (44.52%)	342 (38.64%)	149 (16.84%)
K-means (n=885)	370 (41.81%)	348 (39.32%)	167 (18.87%)

The structure of the data remained constant in the stability analysis, showing that the patterns in the subgroup distribution were reproducible even if random parts of the sample were excluded. Table 1 shows one of the many random partitions that were carried out, all with very similar results. The K-means algorithm also showed a similar pattern in the data, producing roughly the same patient distribution among clusters. Generally, these robustness checks confirmed that the 3-cluster solution and the resulting proportion of patients assigned to each cluster were robust.

Cluster characteristics

The representative features of each cluster are described in Table 2 and Table 3. We report mean values or proportions, together with the significance test p-values and post hoc test results. Some of the reported variables varied across clusters. Age increased significantly from one cluster to the next. The Barthel scale was significantly different for all patient types, showing extensive declines in functional status from cluster to cluster. Moreover, serum albumin values were also significantly different across clusters, indicating poorer nutritional status. The majority of professional-rated variables increased significantly among clusters, as shown by global status, clinical severity and nursing needs. Social needs were the exception in this group of variables, as no significant differences were reported.

Table 2: Patient Characteristics within Clusters

	All the population n=885	Cluster 1 n=400	Cluster 2 n=336	Cluster 3 n=149	ANOVA / Kruskal-Wallis / $\chi 2$ test
1-year mortality risk score (%)	32.15% ±18.04%	16.95% ±6.45%	36.64% ±6.67%	62.84% ±10.68%	KW<0.001*
Age (mean ± s.d.)	83.33 ± 8.37	81.16 ± 8.51	84.32 ± 8.05	86.89 ± 6.97	KW<0.001*
Sex (% Males)	56%	62%	55%	44%	$\chi 2$ <0.001‡
Informal Carer (%)	91%	88%	94%	93%	$\chi 2$: 0.011§
GMA score (mean ± s.d.)	22.81 ± 6.82	22.47 ± 6.39	23.04 ± 7.06	23.24 ± 7.33	AN: 0.379
Barthel scale (mean ± s.d.)	59.95 ± 29.45	82.03 ± 18.16	51.64 ± 17.24	19.40 ± 22.15	KW<0.001*
Albumin/Creatinine Index (mean ± s.d.)	1.49 ± 0.67	1.49 ± 0.66	1.51 ± 0.70	1.46 ± 0.66	KW: 0.792
≤30 mg/g: Normal (n, %)	540 (61%)	242 (61%)	203 (60.42%)	95 (64%)	
30-300 mg/g: Moderate (n, %)	254 (29%)	120 (30%)	94 (28%)	40 (27%)	
≥300mg/g: High (n, %)	91 (10%)	38 (9%)	39 (12%)	14 (9%)	
Creatinine (mg/dL) (mean ± s.d.)	1.52 ± 0.78	1.40 ± 0.59	1.57 ± 0.75	1.72 ± 1.16	KW: 0.002†
Serum albumin (g/dL) (mean ± s.d.)	3.80 ± 0.43	3.90 ± 0.37	3.76 ± 0.41	3.59 ± 0.50	KW<0.001*
Pressure Skin Ulcers (%)	28%	21%	26%	55%	$\chi 2$ <0.001‡
Number of prescriptions (mean ± s.d.)	8.08 ± 3.55	7.82 ± 3.48	8.47 ± 3.75	7.95 ± 3.20	AN: 0.038§
Intake of Opioids (%)	15%	17%	12%	13%	$\chi 2$: 0.168
Intake of Psycholectics (%)	8%	5%	9%	16%	$\chi 2$ <0.001‡
Global Severity (mean ± s.d.)	2.87 ± 0.45	2.63 ± 0.50	2.99 ± 0.09	3.22 ± 0.41	KW<0.001*
Mild (n, %)	4 (<1%)	4 (1%)	-	-	
Moderate (n, %)	144 (16%)	141 (35%)	3 (1%)	-	
Severe (n, %)	704 (80%)	25 (64%)	333 (99%)	116 (78%)	
Very severe (n, %)	33 (4%)			33 (22%)	
Clinical Severity (mean ± s.d.)	3.30 ± 0.68	3.09 ± 0.75	3.43 ± 0.59	3.58 ± 0.55	KW<0.001*
Mild (n, %)	2 (<1%)	2 (1%)	-	-	
Moderate (n, %)	110 (12%)	89 (22%)	17 (5%)	4 (3%)	
Severe (n, %)	393 (44%)	180 (45%)	158 (47%)	55 (37%)	
Very severe (n, %)	380 (43%)	129 (32%)	161 (48%)	90 (60%)	
Nursing Needs (mean ± s.d.)	2.65 ± 0.88	2.23 ± 0.80	2.82 ± 0.77	3.41 ± 0.66	KW<0.001*
Mild (n, %)	76 (9%)	66 (16%)	10 (3%)	-	
Moderate (n, %)	323 (36%)	204 (51%)	105 (31%)	14 (10%)	
Severe (n, %)	318 (36%)	103 (26%)	155 (46%)	60 (40%)	
Very severe (n, %)	168 (19%)	27 (7%)	66 (20%)	75 (50%)	
Social Needs (mean ± s.d.)	1.77 ± 0.70	1.71 ± 0.68	1.82 ± 0.71	1.84 ± 0.73	KW: 0.095
Mild (n, %)	327 (37%)	162 (40%)	116 (35%)	49 (33%)	
Moderate (n, %)	451 (51%)	199 (50%)	173 (51%)	79 (53%)	
Severe (n, %)	90 (10%)	33 (8%)	40 (12%)	17 (11%)	
Very severe (n, %)	17 (2%)	6 (2%)	7 (2%)	4 (3%)	

KW: Kruskal-Wallis test; AN: ANOVA; $\chi 2$: Chi squared test
Missing values that were imputed for analysis (% missing): creatinine (0.23%), albumin/creatinine index (2.15%) clinical severity (3.62%), nursing needs (9.38%), global severity (10.96%), Barthel scale (14.24%), serum albumin (15.71%), social needs (43.62%)
*All clusters are significantly different between them
†Cluster 1 is significantly different to the remaining clusters
‡Cluster 3 is significantly different to the remaining clusters
§Clusters 1 and 2 significantly different, but no other differences in remaining pairwise post-hoc tests

We can therefore say that mortality risk scores increase as age progresses, alongside a deterioration of functional status, nutritional values, clinical severity status and nursing needs status (Table 2). While these trends are common to all patients, certain features inherent to particular clusters were observed:

- **Cluster 1**: risk scores [0% - 26.70%]

  Patients in this cluster showed a lower renal deterioration degree, as measured by creatinine, in comparison to the rest of the clusters. The number of prescriptions and the proportion of patients who had an informal caregiver were significantly lower than those of cluster 2, but there were no significant differences with respect to cluster 3.

- **Cluster 2**: risk scores [26.70% - 50.80%]

  No particular differences were reported with respect to the other clusters, apart from the common differences that relate to age, functional status, nutritional values and the professional-rated variables mentioned above.

- **Cluster 3**: risk scores [50.80% - 100%]

  All patients included in this subgroup had a higher likelihood of dying than of surviving the following year. Regarding specific cluster features, we found a higher proportion of female patients than in other subgroups. Patients included in this subgroup presented a notable increase in the existence of pressure ulcers, with more than 50% of them presenting this problem. Regarding diagnosis, patients in this cluster presented a significantly lower prevalence rate of COPD but higher prevalence rates of cerebrovascular disease and dementia in comparison to those in clusters 1 and 2. In line with the higher prevalence of dementia, a higher intake of psycholectics also was observed.

Despite the many differences described above, some other characteristics remained unchanged across clusters. This was the case for the albumin/creatinine index, which is an early screener for kidney disease, the intake of opioids, and social needs, as highlighted earlier.

The GMA risk score, which was the metric used to select the top 5% of risks for our study, was similar for all clusters. That is, while the mortality risk scores varied among different clusters, the GMA risk scores remained the same.

All other diagnoses apart from COPD, cerebrovascular disease and dementia had similar prevalence rates among the clusters. Regarding the most frequent illness combinations present in our population, no significant differences across clusters were observed (Table 3). The number of coexisting chronic conditions was also equal in all subgroups.

Table 3: Patient Diagnoses within Clusters

	All the population n=885	**Cluster 1 n=400**	**Cluster 2 n=336**	**Cluster 3 n=149**	
Number of comorbidities (mean ± s.d.)	3.65 ± 0.81	3.66 ± 0.82	3.67 ± 0.84	3.61 ± 0.75	KW: 0.928
Prevalence of illnesses					**χ2 test**
Diabetes (n, %)	635 (72%)	289 (72%)	237 (71%)	109 (73%)	p: 0.803
Chronic Renal Failure (n, %)	591 (67%)	257 (64%)	234 (70%)	100 (67%)	p: 0.301
Ischemic Heart Disease (n, %)	444 (50%)	212 (53%)	169 (50%)	63 (42%)	p: 0.082
Heart Failure (n, %)	595 (67%)	267 (67%)	234 (70%)	94 (63%)	p: 0.352
Cerebrovascular Disease (n, %)	283 (32%)	113 (28%)	106 (32%)	64 (43%)	p: 0.004‡
COPD (n, %)	263 (30%)	135 (34%)	102 (30%)	26 (17%)	p: 0.001‡
Asthma (n, %)	190 (21%)	98 (25%)	66 (20%)	26 (17%)	p: 0.118
Dementia (n, %)	166 (19%)	56 (14%)	60 (18%)	50 (34%)	p <0.001‡
Cirrhosis (n, %)	68 (8%)	36 (9%)	26 (8%)	6 (4%)	p :0.151
Most frequent illnesses combinations					
Diabetes + Chronic Renal Failure + Heart Failure (n, %)	67 (8%)	26 (7%)	29 (9%)	12 (8%)	p: 0.537
Diabetes + Chronic Renal Failure + Heart Failure + Ischemic Heart Disease (n, %)	50 (6%)	20 (5%)	19 (6%)	11 (7%)	p: 0.561
Chronic Renal Failure + Heart Failure + Ischemic Heart Disease (n, %)	36 (4%)	16 (4%)	16 (5%)	4 (3%)	p: 0.563
Diabetes + Chronic Renal Failure + Ischemic Heart Disease (n, %)	34 (4%)	16 (4%)	16 (5%)	2 (1%)	p: 0.190
Diabetes + Heart Failure + Ischemic Heart Disease (n, %)	25 (3%)	14 (4%)	5 (1%)	6 (4%)	p: 0.162
Diabetes + Heart Failure + COPD (n, %)	25 (3%)	13 (3%)	8 (2%)	4 (3%)	p: 0.773
Chronic Renal Failure + Heart Failure + COPD	17 (2%)	10 (3%)	6 (2%)	1 (1%)	p: 0.372
Heart Failure + Ischemic Heart Disease + Cerebrovascular Disease (n, %)	18 (2%)	5 (1%)	10 (3%)	3 (2%)	p: 0.255
Diabetes + Heart Failure + Cerebrovascular Disease (n, %)	18 (2%)	7 (2%)	7 (2%)	4 (3%)	p: 0.786
Diabetes + Chronic Renal Failure + Heart Failure + Asthma (n, %)	15 (2%)	10 (3%)	2 (1%)	3 (2%)	p: 0.130

χ2: Chi squared test
‡Cluster 3 is significantly different to the remaining clusters

DISCUSSION

This study demonstrates the existence of clinically distinct subgroups within a population of high-risk patients with multiple chronic conditions, confirming that case management integrated care programs use a uniform treatment strategy for patients who have diverse needs. That is, case management treats heterogeneous populations in a homogeneous way.

The need for a data-based, high-risk patient stratification has been extensively illustrated in the literature, but despite its potential it remains underdeveloped, and only a few studies exist [27]. One of the reasons why this may be the case is that proprietary stratification algorithms such as CRGs are already in place, so healthcare professionals or managers do not see the necessity of using alternative approaches. However, while these algorithms provide considerably better solutions than demographic approaches, they are poor risk adjusters when mortality and other clinical outcomes are considered [21,24]. There is a lack of alignment between the purpose of proprietary algorithms, which aim to stratify patient populations based on estimates of future healthcare resource consumption [13], and the purpose of this study, that is, to stratify patients according to their clinical needs.

Our results support these statements and show how the GMA score, the Spanish equivalent of CRGs, does not vary across subgroups, whereas mortality risk scores do differ from cluster to cluster. GMA does not offer the desired level of granularity to observe clinically relevant subpopulations among high-risk patients, resulting in a homogeneous population from a cost point of view, while subpopulations with different needs remain undetected if alternative risk stratification methods are not introduced.

Given that ready-to-use risk stratification tools are not adequate for the purposes of this study, alternative segmenting methods were explored. Big data techniques and cluster analysis in particular have been proposed for these purposes in the literature when electronic records are available, as in our case [9,16,28]. We showed that cluster analysis is a useful tool for producing risk stratifications, providing valuable information for healthcare reform and robust results that are easy to interpret.

With respect to the variables that were used to stratify our population, only clinically related and demographic variables were used. This approach offers several advantages, emphasizing relevant health priorities that should be addressed and informing the design

of new services or the reform of the existing ones [27]. In contrast, the demand for healthcare services does not always inform areas of clinical concern but of cost concern. Health reforms that arise from using utilization rates for risk stratification may go against the interest of the patient, since the aim of the policy maker may be to reduce costs instead of improve population health [29].

Moreover, if utilization rates are to be used, episodes of care should be comparable [24]. All patients should suffer the same health problem or diagnosis, and all demand episodes should be related to the medical area of interest and equally intense or of the same nature. If the former conditions are fulfilled, the quantity of care provided is appropriate for risk stratification. However, this is hardly ever the case, especially in the case of chronic illnesses, and in our study in particular, patients suffered different illness combinations or types of exacerbations, making utilization episodes incomparable.

A limitation of our study is that the population under study suffered from a specific set of chronic illnesses that may not be the same in other settings. In addition, patients only qualified for enrollment if they suffered from three or more chronic illnesses. Other programs may require only two chronic illnesses for enrollment. This may impact the generalizability of our study results. Moreover, we tested the cluster stability and robustness internally rather than externally. As a final limitation, we specified patient clusters using our own risk score estimations. However, different risk scoring models are likely to be used in other environments. We encourage others to reproduce our analyses and estimate risk scores for each context.

One study by Vuik et Al. [30] stratified a high-risk patient population using cluster analysis. However, that study grouped patients according to their utilization patterns and not their clinical risk scores or needs. Moreover, cancer patients were included while we did not include this type of patient. Four main subgroups among which care usage had significant variation were identified. Low et Al. [31] also provided a risk stratification using cluster analysis, using utilization data to group patients. Their study was not restricted to high-risk patients and included all types of adult patients in the analysis, without making distinctions in terms of their risk category or clinical profile. Five clusters were found.

Other studies that segment patient populations are available in the literature, although they used expert criteria to produce the resulting subgroups instead of data-driven

approaches [17,32]. Lynn et al. describe three end-of-life subgroups for frail, high-risk patients, which is in line with our results.

Tailoring integrated care services

To date, all patients included in this study have been treated under the same case management strategy. Nevertheless, three distinct subgroups with different characteristics were identified for which care programs should be tailored.

We proposed a different care strategy for each type of patient so that treatment can adequately meet patient needs. These strategies were based on a literature review and supported by expert consultation with healthcare authorities from the region who have extensive experience with the integrated care program under study.

Patients included in cluster 1, whose risk status was the lowest of all subgroups and who had moderate functional status, severity status and nursing needs, could benefit from their current case management program. This program includes a reference specialist team that keeps patient follow-up, self-care education and support, a link nurse that is available 24 h by phone, and most importantly, direct hospital admission without passing through emergency services and a day hospital unit. At-home services are also available in some cases. All professionals develop personalized care plans that focus on avoiding exacerbations and sustaining function.

As health starts to decline in combination with a worsening functional status and increasing nursing needs, patients become increasingly dependent on a 3rd person, and transitions from home to the hospital can be complicated. Patients in cluster 2 would benefit from home-based programs that focus on improving quality of life and averting unnecessary hospitalizations or readmissions [32]. Mobile integrated care programs should be implemented for these purposes [33]. Nursing services, together with caregiver training and support, play an important role.

Those included in cluster 3, with the highest mortality risk scores, are very likely to die in the near future. Continuing functional declines, together with worsening clinical severity and other characteristics such as increases in pressure skin ulcers or higher rates of dementia and cerebrovascular disease, are indicators of the short survival prospect of patients included in this subgroup. Healthcare services should be directed towards end-of-life care, including hospices, or home-based palliative care services that shift attention from curative efforts to quality of life improvements [34].

A risk score estimation tool has been created for use in clinical practice to estimate patient risk scores [23]. This tool consists of a nomogram, which is a graphical calculation tool that synthesizes logit model results in a graph that is filled in by healthcare professionals and provides individual risk scores for each patient without the need for computers or software. Risk scores, in combination with the results of this study, can be used in clinical practice for patient classification purposes. Risk score calculation and subsequent patient classification should be performed at patient enrollment in the program but also at regular intervals or if healthcare professionals see it as necessary. This would allow close monitoring of each patient situation, providing valuable information that can assist treatment strategy decisions.

This study is a key part of the design of alternatives to case management care programs. By stratifying the population into differentiated subpopulations, we identified relevant patient types and their needs. The description of the characteristics of each patient type can guide the development of these new services. Moreover, study results provide valuable information for healthcare professionals in relation to the development of each patient's condition and can assist treatment strategy decisions.

The extent to which patient outcomes such as mortality rates or quality of life improve will be determined in future research when alternative programs are implemented and their performance measured. The efficiency of new care strategies also needs to be measured in future research through cost-effectiveness analyses.

CONCLUSIONS

This study produced a restratification for a population of high-risk multimorbid patients who are currently included in a case management integrated care program. We showed that the high-risk population had heterogeneous needs but that all patients received the same treatment. Risk stratification was performed using cluster analysis. The characteristics of each cluster were presented, outlining the specific needs that should be addressed in healthcare reform. We suggested alternatives to case management services that can make meaningful contributions to health outcomes, moving away from low-value care.

REFERENCES

1. Ham C. The ten characteristics of the high-performing chronic care system. Health Economics, Policy and Law, 5, pp71-90 (2010) doi.org/10.1017/S1744133109990120
2. Lenzi J, Avaldi VM, Rucci P et al. Burden of multimorbidity in relation to age, gender and immigrant status: a cross-sectional study based on administrative data. BMJ Open; 6: e012812. (2016) doi.org/10.1136/bmjopen-2016-012812
3. Navickas R, Petric VK, Feigl AB et al. Multimorbidity: What Do We Know? What Should We Do? Journal of Comorbidity, 4–11. (2016) doi:10.15256/joc.2016.6.72.
4. Sinnige J, Braspenning J, Schellevis F, et al. The Prevalence of Disease Clusters in Older Adults with Multiple Chronic Diseases – A Systematic Literature Review. PLoS ONE 8(11): e79641. (2013) doi:10.1371/journal.pone.0079641
5. Orueta, J.F., Nuño-Solinis, R., Mateos, M. et al. Predictive risk modelling in the Spanish population: a cross-sectional study. BMC Health Serv Res 13, 269 (2013) doi.org/10.1186/1472-6963-13-269
6. Wagner, E., Austin, B., & Von Korff, M. Organizing Care for Patients with Chronic Illness. The Milbank Quarterly 74, no. 4: 511-44. (1996) doi:10.2307/3350391.
7. Ham C., Curry N., Clinical and service integration: The route to improved outcomes. The Kings Fund, ISBN978 1 85717 605 6 (2010) www.kingsfund.org.uk/publications/clinical-and-service-integration. Accessed July 2020
8. Satylganova A., Integrated care models: an overview. World Health Organisation (2016) www.euro.who.int/en/health-topics/Health-systems/health-services-delivery/publications/2016/integrated-care-models-an-overview-2016. Accessed July 2020
9. Vuik S.I., Mayer E.K., Darzi A., Patient Segmentation Analysis Offers Significant Benefits For Integrated Care And Support. Health Affairs 35:5, 769-775 (2016) doi: 10.1377/hlthaff.2015.1311
10. Orueta JF, García Alvarez A, Aurrekoetxea JJ, et al. FINGER (Forming and Identifying New Groups of Expected Risks): developing and validating a new predictive model to identify patients with high healthcare cost and at risk of admission. BMJ Open; 8:e019830. (2018) doi:10.1136/bmjopen-2017-019830
11. Juhnke, C., Bethge S., Mühlbacher A.C., A Review on Methods of Risk Adjustment and their Use in Integrated Healthcare Systems. International Journal of Integrated Care, 16(4): 4, pp. 1–18, (2016) doi: http://dx.doi.org/10.5334/ijic.2500
12. Zhou Y.Y., Wong W., Li J., Improving Care for Older Adults: A Model to Segment the Senior Population. Perm J; 18(3):18-21 (2014) doi.org/10.7812/TPP/14-005
13. Hughes J., Averill R., Eisenhandler J. et al. Clinical Risk Groups (CRGs): A Classification System for Risk-Adjusted Capitation-Based Payment and Health Care Management. Medical Care; 42(1):81-90, (2004) doi:10.1097/01.mlr.0000102367.93252.70
14. Lanzeta I, Mar J, Arrospide A. Cost-utility analysis of an integrated care model for multimorbid patients based on a clinical trial; Gac Sanit, 2016; 30(5): 352-358. doi.org/10.1016/j.gaceta.2016.05.002.
15. Jean-Baptiste D., O'Malley A., Shah T., Population Segmentation and Targeting of Health Care Resources: Findings from a Literature Review. Mathematica Policy Research, Working Paper 58 (2017)

16. Johnson T.L., Rinehart D.J., Durfee J., et al. For Many Patients Who Use Large Amounts of Health Care Services, the Need Is Intense Yet Temporary Health Affairs 34:8, 1312-1319 (2015) doi: 10.1377/hlthaff.2014.1186
17. Joynt K.E., Figueroa J.F., Beaulieu N., et al. Segmenting high-cost Medicare patients into potentially actionable cohorts, Healthcare, Volume 5, Issues 1–2, Pages 62-67, (2017) doi.org/10.1016/j.hjdsi.2016.11.002.
18. Verkerk E.W, Tanke M.A.C, Kool R.B, et Al. Limit, lean or listen? A typology of low-value care that gives direction in de-implementation, International Journal for Quality in Health Care, Volume 30, Issue 9, November 2018, Pages 736–739, https://doi.org/10.1093/intqhc/mzy100
19. Ibarrola C., Sanz A., Iribarren I., et al. Estrategia Navarra De Atención Integrada A Pacientes Crónicos Y Pluripatológicos (2016) Available at https://www.navarra.es/home_es/Temas/Portal+de+la+Salud/Ciudadania/Nuevo+Modelo+asistencial/Estrategia+navarra+de+atencion+integrada+a+pacientes+cronicos+y+pluripatologicos/ Accessed July 2020
20. Monterde, D., Vela, E., Clèries, M et al. Los grupos de morbilidad ajustados: nuevo agrupador de morbilidad poblacional de utilidad en el ámbito de la atención primaria. Atención Primaria, 48(10), 674-682. (2016) doi: 10.1016/j.aprim.2016.06.003
21. Dueñas-Espín I, Vela E, Pauws S, et al. Proposals for enhanced health risk assessment and stratification in an integrated care scenario, BMJ Open, 6:e010301. (2016) doi:10.1136/bmjopen-2015-010301
22. Basagaña X., Barrera-Gómez J., Benet M., et al., A Framework for Multiple Imputation in Cluster Analysis, American Journal of Epidemiology, Volume 177, Issue 7, 1, Pages 718–725, (2013) doi.org/10.1093/aje/kws289
23. Bretos-Azcona PE, Ibarrola Guillén C, Sánchez-Iriso E, et Al. Multisystem chronic illness prognostication in non-oncologic integrated care [published online ahead of print, 2020 Jun 24]. BMJ Support Palliat Care. 2020;bmjspcare-2019-002055. doi:10.1136/bmjspcare-2019-002055
24. Iezzoni L.I., Risk Adjustment for Measuring Health Care Outcomes, 4th edition. Health Administration Press, Chicago, Illinois ISBN-13: 978-1567934373 (2013)
25. Everitt B.S., Landau S., Leese M., et al., Cluster Analysis, 5th Edition, John Wiley & Sons, Ltd. ISBN: 978-0-470-74991-3 (2011)
26. Rousseeuw P.J., Silhouettes: A graphical aid to the interpretation and validation of cluster analysis, Journal of Computational and Applied Mathematics, Volume 20,Pages 53-65, (1987) doi.org/10.1016/0377-0427(87)90125-7.
27. Chong, J.L., Lim, K.K., Matchar, D.B. Population segmentation based on healthcare needs: a systematic review. Syst Rev 8, 202 (2019). doi.org/10.1186/s13643-019-1105-6
28. Bates D.W., Saria S., Ohno-Machado L., et al. Big Data in Health Care: Using Analytics to Identify And Manage High-Risk And High-Cost Patients Health Affairs 2014 33:7, 1123-1131(2014) doi: 10.1377/hlthaff.2014.0041
29. Wagner EH. Chronic disease management: What will it take to improve care for chronic illness? Effective Clinical Practice; 1:2-4. (1998)
30. Vuik SI, Mayer E, Darzi A. Enhancing risk stratification for use in integrated care: a cluster analysis of high-risk patients in a retrospective cohort study. BMJ Open; 6: e012903. (2016) doi:10.1136/bmjopen-2016-012903
31. Low LL, Yan S, Kwan YH, et al. Assessing the validity of a data driven segmentation approach: a 4 year longitudinal study of healthcare utilization and mortality. PLoS One.; 13(4):e0195243(2018) doi.org/10.1371/journal.pone.0195243

32. Lynn J, Straube BM, Bell KM, et al. Using population segmentation to provide better health care for all: the "Bridges to Health" model. Milbank Q.; 85(2):185–212. (2007) doi:10.1111/j.1468-0009.2007.00483.x
33. Clarke J.L, Bourn S., Skoufalos A., et al. An Innovative Approach to Health Care Delivery for Patients with Chronic Conditions; Population Health Management 20:1, 23-30 (2017) doi:10.1089/pop.2016.0076
34. Boyd, K., Murray, S. Recognising and managing key transitions in end of life care. BMJ: British Medical Journal, 341(7774), 649-652 (2010) doi.org/10.1136/bmj.c4863

CHAPTER 3

SURVIVAL PATTERNS OF NON-ONCOLOGIC HIGH-RISK PATIENTS IN INTEGRATED CARE

ABSTRACT

Background

Delivering palliative care for patients depends on the ability to predict the timing of death. Given the difficulty to predict survival in non-oncologic chronic patients, these have limited access to end of life care. The objective of this study is to produce time to death estimates and to find prognostic variables in a population of high-risk, non-oncological patients that are treated in an integrated care programme.

Methods

We distinguish three patient subgroups in our analyses according to their end of life trajectory: *Terminal Illness, Organ Failure and Frailty*. All high-risk multiple chronic conditions patients from the Spanish region of Navarra were considered in this study (n=885). Kaplan-Meier estimates were completed for each patient subgroup. A Cox proportional hazard model was constructed to evaluate the relationship of every patient characteristic with respect to survival. The proportional hazards (PH) assumption of the model was tested on the basis of Schoenfeld residuals. Additional parametric models to predict survival beyond study period limits were performed.

Results

Median Kaplan-Meier estimates were 734 days and 244 days for *Organ Failure* and *Frail* patients. More than 50% of *Terminal Illness* patients alive by the end of the follow-up period. Variables found to be independently related to survival and fulfilled the proportional hazards assumption in the Cox multivariate model were age, the presence of pressure skin ulcers, prevalence of heart failure, and global severity. Weibull parametric model results were similar to those shown in Cox PH regression results, and median survival estimates were 1805.83 days, 921.02 days and 324.97 days for *Terminal Illness, Organ Failure* and *Frail* types of patients respectively.

Conclusion

This study provides empirical evidence on time to death estimates in a population of non-cancer multimorbid patients in an integrated care setting, helping to organise end of life care for these patients.

INTRODUCTION

Most individuals in economically advanced countries develop chronic illnesses at the end of their lives [1]. As a consequence, elderly patients tend to live in an overall state of functional and clinical decline that shapes their last years of life. The situation of these patients worsens gradually with time producing severe disability, exacerbations, and finally causing death [2]. This is especially true in patients that suffer multiple chronic conditions (MCC) at the same time, who account for high levels of mortality and resource utilisation at the end of life [3, 4]. The prevalence of chronic illness during end of life is high [5], and in Spain the most common causes of death are now chronic conditions rather than acute episodes of illness [6, 7].

Delivering better care during the end of life for these patients depends in many cases on being vigilant to the possibility of needing palliative care [1]. Therefore, an early identification of potential palliative care patients is central to end of life care strategies [8].

A typical approach to proceed with early identification is to classify high-risk patients in terms of the different illness trajectories before death that exist [2, 9]. Trajectories aim to describe the course of decline in terms of length and rate of functional decline [9-11]. This informs clinicians about patterns of probable needs, helping to plan, set priorities and deliver appropriate care to each type of patient [12].

The literature identifies three main, distinct illness trajectories at the end of life [8, 10, 11, 13] The first of these is the "*Terminal Illness*" trajectory, which describes a group of patients that maintain functional status for a substantial period of time, followed by an abrupt and severe terminal phase of decline in performance status in the final weeks preceding death [2, 8-10, 12, 13]. The "*Organ Failure*" trajectory is characterised by a gradual decline in functional status, that erodes when patients experience intermittent episodes of acute deterioration usually related to underlying disease evolution. While every exacerbation is potentially deadly, patients usually recover from several crises but never to baseline showing an unstable pattern of decline [8-13]. Finally, patients who escape the previous end of life patterns may also experience the "*Frail*" trajectory. It consists of a steady downward path of progressive decline that advances in a very slow and gradual way until death, rather than suffering a terminal phase or acute

exacerbations. Patients start off with a low baseline of cognitive or physical functioning, and continue to decline almost imperceptibly until death [2, 8-12].

In the past, an additional "*Sudden Death*" trajectory was the most commonly observed end of life trajectory. The majority of deaths were caused in acute episodes of need, progressing from normal function to death without advance notice [9-11]. However, today this trajectory is not common, particularly in chronically ill patients, and people usually follow one of the other three trajectories [12].

The timing of palliative care usually depends on the predictable nature of the disease trajectory in their last months of life. This facilitates the identification of patients that will die soon and therefore need palliative care. This is the reason why palliative services are predominantly used by cancer patients [14], for which survival estimates are generally available.

However, non-oncologic patients have limited access to palliative care given the difficulty to predict survival in this population [2, 10, 14], especially in multiple chronic conditions patients that suffer several conditions at the same time [15]. Uncertainty in time to death in the different trajectories complicates the decision regarding when to start palliative care [1, 16].

Objectives

The objective of this study is to estimate time to death for each of the three main end of life trajectory groups that have been described above in a population of high-risk, non-oncological patients that are treated in an integrated care programme. We also aim to determine whether survival estimates for non-cancer patients are predictable and to find prognostic variables that predict time to death.

METHODS

The healthcare system of the Spanish region of Navarra covers a population of approximately 650,000 individuals. In April 2016, an integrated care programme was put in place to treat non-oncologic MCC patients [17]. The programme promotes self-management for low-risk patients, and introduces elements from the Chronic Care Model as patient risk increases. For those patients that belong to the top of the programmes' risk pyramid, case management services such as a reference specialist

team, 24h nurse phone contact, hospital admission without passing through emergency services or a day hospital are available.

The high-risk population included in case management is considered to be the top 5% risks according to the GMA score (Grupos de Morbilidad Ajustados), which is the Spanish equivalent to the Clinical Risk Groups (CRGs) stratification tool [18]. To qualify for enrolment patients must suffer at least three chronic illnesses from a list that includes Heart Failure, Dementia, Ischemic Heart Disease, Cerebrovascular Disease, Diabetes, Chronic Obstructive Pulmonary Disease, Asthma, Chronic Renal Failure and Cirrhosis. This selection of illnesses is defined by the programme's framework.

In a recent risk stratification of the high-risk MCC population, three heterogeneous subgroups have been identified using a risk score based on baseline functional status, patient severity and other end-of-life features such as pressure skin ulcers and renal deterioration. The resulting risk strata represented the top 0.84%, the top 1.90-2.26% and the top 2.26-5% risks of the case management model respectively [19]. We present baseline patient characteristics at enrolment for each patient subgroup in Table 1.

All high-risk MCC patients from the region enrolled since programme implementation in April 2016 until August 2018 (n=885) were included into analysis and followed-up until January 2020. The mean follow-up length was 636 (1.74 years) and median follow-up was 667 days (1.83 years).

At enrolment, patients are subject to a comprehensive assessment of their situation. A set of baseline variables are collected at this point in time, including age, sex, functional status (Barthel scale), nutritional status (serum albumin level), renal deterioration status (creatinine level), presence of pressure skin ulcers, and global status, which is a variable indicating patient severity. Global status score is produced by the case management team responsible for each patient (doctors, nurses and social services workers) based on expert professional criteria. Other professional-rated scales informing of particular areas of interest such as clinical severity, nursing needs or social needs are available. Prevalence of selected non-cancer illnesses is also recorded. The date of inclusion into the programme, as well as the date of death are available for survival analysis.

For those patient characteristics that had missing values, multiple imputation procedures were implemented using chained equations under a missing at random (MAR)

assumption. This ensures that no observations are lost in analyses while avoiding biased results.

Survival analyses

For each patient, time to death since enrolment into the programme, or time to study end if patients were censored were computed. Kaplan-Meier non-parametric analyses were performed to estimate mean and median survival time for all three high-risk patient subgroups. Survival curves are presented in Figure 1, together with the number of patients at risk. The logrank test was used to determine whether the survival curves were significantly different from each other.

Univariate Cox proportional hazard regression analyses were constructed to evaluate the effect of every patient characteristic with respect to survival. Variables with p-values lower than 0.25 were further assessed in multivariate models. A final Cox regression model was built using backwards stepwise methods, with a significance value of $p \leq 0.05$. Cox predicted curves for the final model were plotted against Kaplan-Meier curves to discern whether there is a substantial difference between both sets of curves.

The proportional hazards (PH) assumption of the model was tested on the basis of Schoenfeld residuals for each of the variables that were included in the final model, and for the model as a whole. The PH was further tested visually by comparing the graphs of the baseline survival function for each of the patient subgroups under analysis, which are shown in Supplementary File 1.

Once the final Cox regression model was fitted and the PH assumption tested, significant variables were used to build additional parametric models to expand survival estimates. These models allow predicting survival beyond study period limits and can shed light about how curves would have looked like if no censoring would have occurred.

To decide which distribution among Weibull, Gompertz, lognormal, loglogistic, and generalized gamma would fit best into the parametric model, all models were completed and the one with the lowest Akaike information criterion (AIC) was chosen. Expanded survival curves were plotted together with Kaplan-Meier curves for reference. All analyses were conducted in STATA 15.1 software.

RESULTS

Patient characteristics

Table 1 shows the baseline characteristics of all distinct patient types for which survival patterns are analysed. A gradual deterioration in health condition is observed across patient subgroups. Of special relevance is the sharp decline in baseline functional status that is experienced from one group to another. Moreover, patients show a worsening in global, clinical and nursing severity scores. Nutritional (serum albumin) and renal (creatinine) status also deteriorate as patients age.

Patients at the top of the risk pyramid have some additional characteristics that differentiate them, including higher rates of pressure skin ulcers and higher prevalence in dementia and cerebrovascular disease. COPD rates are higher in less severe groups of patients. Not all features differ among groups, for example the number of comorbidities or social needs are similar for all patient types.

These characteristics meet the definitions that are made in the literature for '*Terminal Illness'*, '*Organ Failure'* and '*Frail'* patient subgroups. Therefore this study constitutes a good opportunity to provide time to death estimates and gain an understanding of which parameters have a significant influence towards survival.

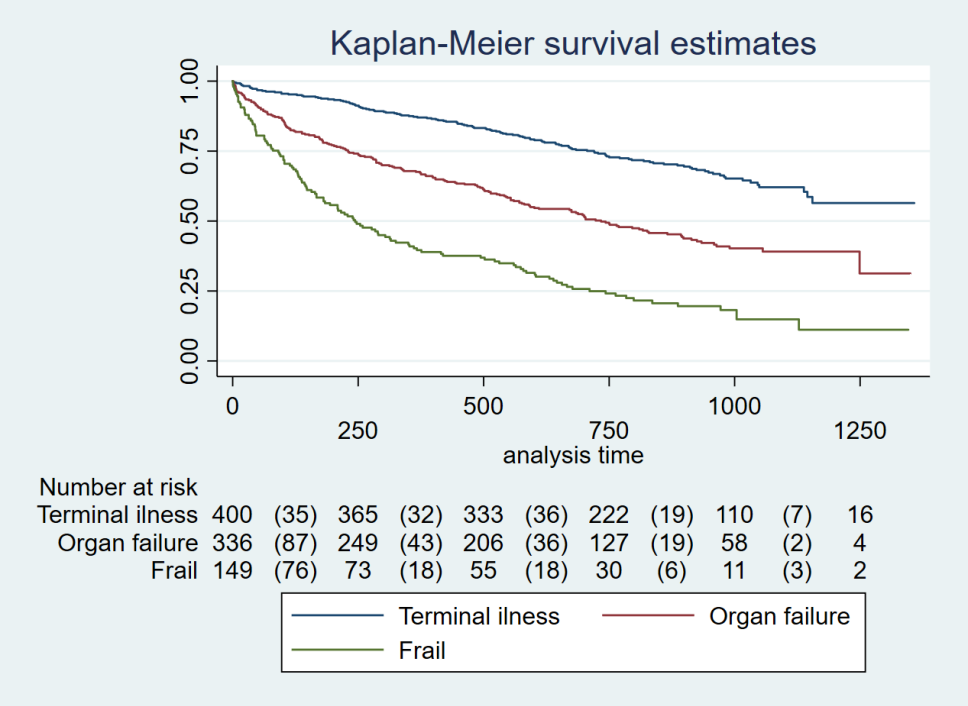
Table 1: Baseline Patient Characteristics and Univariate Cox Proportional Hazards regression results	All high-risk patients (n=885)	Terminal illness (n=400)	Organ failure (n=336)	Frail (n=149)	Univariate Cox PH regression Hazard Rates	Univariate Cox PH regression p-values
Age (mean ± s.d.)	83.33 ± 8.37	81.16 ± 8.51	84.32 ± 8.05	86.89 ± 6.97	1.0511	0.000
Sex (% Males)	56%	62%	55%	44%	0.9727	0.774
Barthel scale (mean ± s.d.)	59.95 ± 29.45	82.03 ± 18.16	51.64 ± 17.24	19.40 ± 22.15	0.9851	0.000
Creatinine (mg/dL) (mean ± s.d.)	1.52 ± 0.78	1.40 ± 0.59	1.57 ± 0.75	1.72 ± 1.16	1.1258	0.031
Serum albumin (g/dL) (mean ± s.d.)	3.80 ± 0.43	3.90 ± 0.37	3.76 ± 0.41	3.59 ± 0.50	0.6326	0.000
Pressure Skin Ulcers (%)	28%	21%	26%	55%	1.6124	0.000
Global Severity						
Mild (n, %)	4 (<1%)	4 (1%)	-	-	-	-
Moderate (n, %)	144 (16%)	141 (35%)	3 (1%)	-	-	-
Severe (n, %)	704 (80%)	255 (64%)	333 (99%)	116 (78%)	2.4920	0.000
Very severe (n, %)	33 (4%)	-	-	33 (22%)	8.7668	0.000
Clinical Severity						
Mild (n, %)	2 (<1%)	2 (1%)	-	-	-	-
Moderate (n, %)	110 (12%)	89 (22%)	17 (5%)	4 (3%)	-	-
Severe (n, %)	393 (44%)	180 (45%)	158 (47%)	55 (37%)	0.9597	0.680
Very severe (n, %)	380 (43%)	129 (32%)	161 (48%)	90 (60%)	3.9560	0.000
Nursing Needs						
Mild (n, %)	76 (9%)	66 (16%)	10 (3%)	-	-	-
Moderate (n, %)	323 (36%)	204 (51%)	105 (31%)	14 (10%)	1.3892	0.175
Severe (n, %)	318 (36%)	103 (26%)	155 (46%)	60 (40%)	2.0901	0.001
Very severe (n, %)	168 (19%)	27 (7%)	66 (20%)	75 (50%)	3.1369	0.000
Social Needs						
Mild (n, %)	327 (37%)	162 (40%)	116 (35%)	49 (33%)	-	-
Moderate (n, %)	451 (51%)	199 (50%)	173 (51%)	79 (53%)	1.0270	0.843
Severe (n, %)	90 (10%)	33 (8%)	40 (12%)	17 (11%)	0.9916	0.959
Very severe (n, %)	17 (2%)	6 (2%)	7 (2%)	4 (3%)	1.5160	0.106
Number of comorbidities (mean ± s.d.)	3.65 ± 0.81	3.66 ± 0.82	3.67 ± 0.84	3.61 ± 0.75	0.9971	0.960
Diabetes (n, %)	635 (72%)	289 (72%)	237 (71%)	109 (73%)	0.9365	0.534
Chronic Renal Failure (n, %)	591 (67%)	257 (64%)	234 (70%)	100 (67%)	0.9022	0.307
Ischemic Heart Disease (n, %)	444 (50%)	212 (53%)	169 (50%)	63 (42%)	0.8413	0.072
Heart Failure (n, %)	595 (67%)	267 (67%)	234 (70%)	94 (63%)	1.3215	0.009
Cerebrovascular Disease (n, %)	283 (32%)	113 (28%)	106 (32%)	64 (43%)	1.1647	0.129
COPD (n, %)	263 (30%)	135 (34%)	102 (30%)	26 (17%)	0.8455	0.118
Asthma (n, %)	190 (21%)	98 (25%)	66 (20%)	26 (17%)	0.9447	0.627
Dementia (n, %)	166 (19%)	56 (14%)	60 (18%)	50 (34%)	1.4336	0.002
Cirrhosis (n, %)	68 (8%)	36 (9%)	26 (8%)	6 (4%)	0.6456	0.034

COPD: Chronic Obstructive Pulmonary Disease; PH: Proportional Hazards Missing values that were imputed for analysis (% missing): creatinine (0.23%), albumin/creatinine index (2.15%) clinical severity (3.62%), nursing needs (9.38%), global severity (10.96%), Barthel scale (14.24%), serum albumin (15.71%), social needs (43.62%)

Survival results

Survival patterns are clearly distinguishable in Figure 1 and are statistically different from one another (logrank test p=0.000). *Frail* patients die at a fastest rate, with the majority of deaths occurring during the first months since enrolment and a small share of patients surviving until the end of study follow-up. In contrast, *Terminal Illness* patients have much better prospects. In fact, less than 50% of those individuals died during the study period. In between both types of patients, *Organ Failure* patients also experienced a rapid decay with high mortality rates, although not as high as for *Frail* patients.

Figure 1. Kaplan-Meier survival estimates



Kaplan-Meier estimates were produced for each patient category, and are presented in Table 2. Median survival estimates were 734 days (2.01 years) and 244 days (0.67 years) days for *Organ Failure* and *Frail* patients. In the case of *Terminal Illness*, there were more than 50% of patients alive by the end of the follow-up period. For this reason a median survival estimate cannot be produced. Mean survival estimates were 1036.96 days (2.84 years), 756.36 days (2.07 years) and 445.10 days (1.22 years) for *Terminal Illness, Organ Failure* and *Frail* patients respectively.

Note that only restricted means and medians can be produced at this stage, and therefore Kaplan-Meier results are underestimated. The reason for this is that survival curves were not fully observed during study time. Further analyses that extend survival curves are presented below.

To understand which variables might have a significant influence in survival, a Cox proportional hazards model was implemented. Variables that showed a significant influence in univariate analyses are provided in Table 1. These include age, creatinine, serum albumin, pressure skin ulcers, global severity, nursing severity and some diagnoses such as heart failure, cirrhosis and dementia. Sex did not have a significant effect on survival, in a similar way to social needs and number of comorbidities. Prevalence of diabetes, chronic renal failure, ischemic heart disease, cerebrovascular disease, COPD and asthma were not significant either. Clinical severity was only significant at the very severe level.

With respect to the Cox multivariate model, variables found to be independently related to survival were age, the barthel scale, the presence of pressure skin ulcers, prevalence of heart failure, and global severity. Patients have higher hazard rates as they age, suffer from heart failure, develop pressure skin ulcers and their global severity increases. Functional status was also found to be a statistically significant variable, showing that lower baseline functional status lead to lower survival.

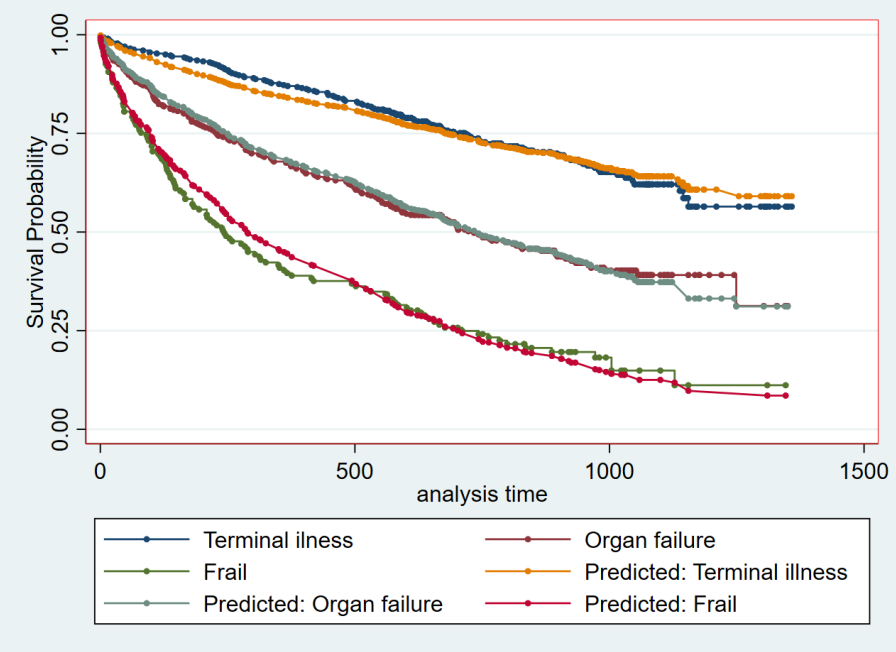
Table 2: Time to death estimates	**Median survival (days)**	**Mean survival (days)**
Kaplan Meier (restricted)		
Terminal illness	-	1036.96
Organ failure	734	756.36
Frail	244	445.10
All high-risk patients	941	832.04
Weibull (extended)		
Terminal illness	1805.83	2543.25
Organ failure	921.02	1773.30
Frail	324.97	626.67
All high-risk patients	1381.52	2489.41

More than 50% of Terminal Illness patients were alive by the end of the follow-up period. For this reason a Kaplan-Meier median survival estimate cannot be produced.

However the individual Schoenfeld residual tests showed that the barthel scale and age rejected the null hypothesis of proportional hazards. The global test was also rejected. As a consequence we fitted another model without the barthel scale. This model did satisfy all the PH conditions. Both models are presented in Table 3, and baseline survival plots are available in Supplementary file 1. We also present in Figure 2 the predicted Cox survival curves, which have been plotted together with Kaplan-Meier curves for reference.

	Model 1				Model 2 (without the Barthel scale)			
Table 3: Multivariate Cox PH regression analyses	**Hazard rate**	**p-value**	**95% CI**	**Schoenfeld residuals p-values**	**Hazard rate**	**p-value**	**95% CI**	**Schoenfeld residuals p-values**
Age	1.030	0.000	1.016 - 1.044	0.030	1.043	0.000	1.029 - 1.057	0.222
Barthel scale	0.990	0.000	0.985- 0.994	0.003	-	-	-	-
Pressure skin ulcers	1.269	0.024	1.033 - 1.560	0.427	1.484	0.000	1.217 - 2.809	0.900
Heart failure	1.388	0.003	1.120 - 1.720	0.831	1.292	0.017	1.047 - 1.595	0.621
Severe global severity	2.109	0.000	1.489 - 2.988	0.335	2.275	0.000	1.612 - 3.212	0.545
Very severe global severity	5.646	0.000	3.493 - 9.127	0.342	6.672	0.000	4.112 - 10.826	0.535
Global Schoenfeld residuals PH test				0.071				0.794

Figure 2. Cox proportional hazards predicted survival curves.



Predicted survival

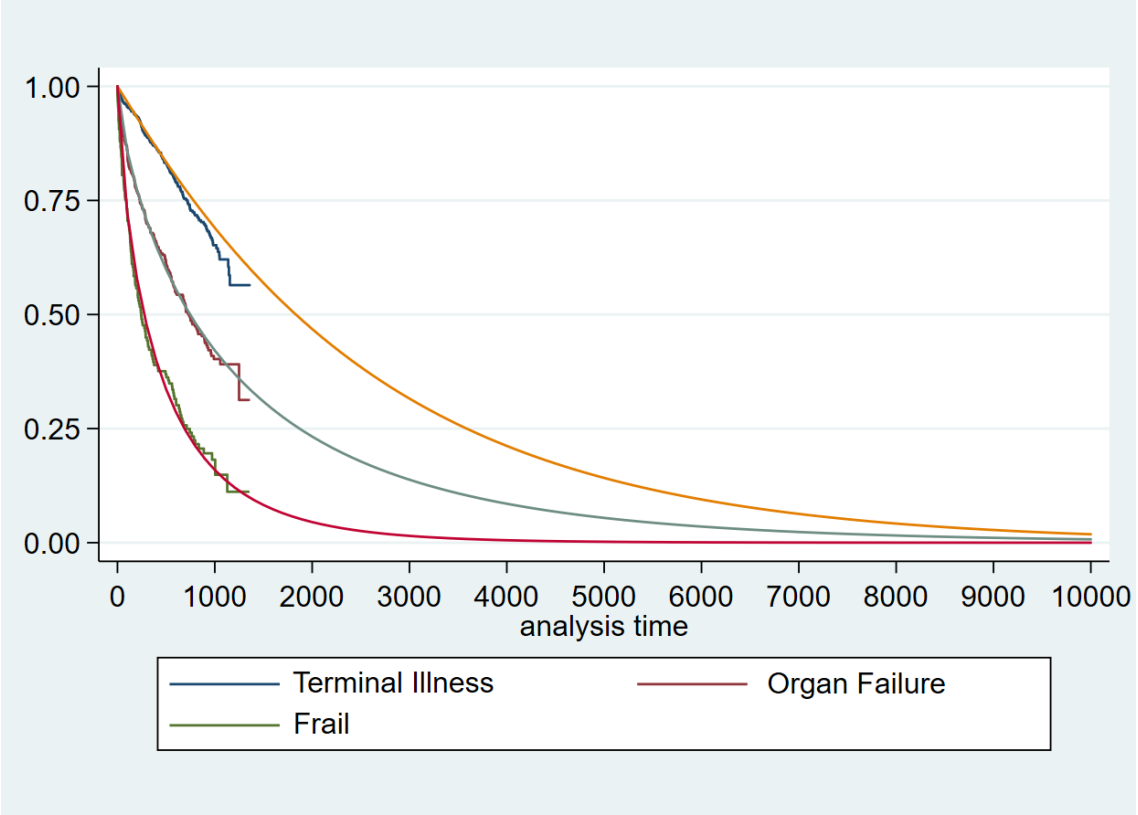
Now that the final Cox PH regression model has been specified, statistically significant variables have been identified and can be used to fit parametric models that can extend survival estimations to the future. The Akaike information criterion (AIC) of several possible distributions (Weibull, Generalised Gamma, Gompertz, Loglogistic, Lognormal) was used to decide which parametric curves best fit to the data. The option with the lowest AIC is the Weibull distribution, and therefore was the chosen option (Table 4).

The Weibull regression results are similar to those shown in Cox PH regression results, yet some additional information can be drawn from the ancillary parameter p. This parameter informs of whether the hazard ratio increases (p>1) or decreases (p<1) over time. Taking all subgroups into account, the hazard decreases with time. However this is driven by *Organ Failure* (p=0.755) and *Frail* subgroups (p=0.754) , that presented a higher proportion of deaths in the first months since programme enrolment as shown in Kaplan-Meier curves. *Terminal Illness* has a p parameter higher than 1 (p=1.032), and therefore we can expect the mortality rate to increase as time goes by.

Table 4: Weibull regression	**Hazard rate**	**p-value**	**95% Conf. Interval**
Age	1.043	0.000	1.029 - 1.057
Pressure skin ulcers	1.486	0.000	1.219 - 1.811
Heart failure	1.291	0.018	1.046- 1.593
Severe global severity	2.260	0.000	1.602 - 3.187
Very severe global severity	6.526	0.000	4.033 - 10.560
Constant	<0.001	0.000	<0.001 – 0.0001
ancillary parameter p	0.796		0.730 - 0.866
1/p	1.257		1.155 - 1.368
ancillary parameter p: Terminal illness	1.032		0.879 - 1.212
ancillary parameter p: organ failure	0.755		0 .662 - 0.860
ancillary parameter p: frail	0.754		0.648 - 0.877

Predicted survival curves are provided in Figure 3. Extended median estimates are 1805.83 days (4.94 years), 921.02 days (2.52 years) and 324.97 days (0.89 years) for *Terminal Illness, Organ Failure* and *Frail* types of patients. Mean survival estimates are 2543.25 days (6.96 years), 1773.30 days (4.86 years) and 626.67 days (1.72 years) for each group.

Figure 3. Weibull regression predicted survival curves.



DISCUSSION

This study provides empirical evidence on time to death estimates in a population of non-cancer, multiple chronic conditions patients treated in an integrated care setting. Three subgroups representing the main end of life illness trajectories are analysed. In addition, prognostic variables have also been identified, and survival curves extrapolations have been produced.

The difficulty to predict time to death in non-cancer patients is one of the main reasons for delaying palliative care in these patients [6, 14]. In contrast, cancer patients have well defined prognoses and use these services intensively [20, 21]. This difference in the use of palliative care can be explained since a reliable estimation on the timing of death is crucial to decide when to transition from curative efforts towards end of life care [1, 2, 22].

While the majority of patients in hospices meet short term predictable death criteria due to their oncologic diagnoses, this model of end of life care may not fit the different needs of non-cancer patients. Having a gradual decline or more prolonged needs does not mean that these are not as demanding as those of cancer patients [12]. In other

words, non-cancer patients should not be discriminated for end of life services admission due to their longer survival prognosis. Instead, they should be considered on the basis of their needs.

The findings presented in this study show not only that survival in non-cancer patients is predictable, but that such predictions are higher to the 6-month survival requirement to qualify for palliative care in some settings like Medicare and Medicaid [21, 23]. Survival prognostic variables and estimations will help to overcome the obstacles in incorporating non-cancer patients to palliative care services [24]. The main illness trajectory to which they belong should also be considered when making decisions, as each type of patients will require a different end of life approach depending on their survival estimates.

The prognostic variables that were identified match what is reported in the literature, with some key differences. Dementia, which is believed to characterise end of life and survival in frail patients [16], was statistically significant only at the univariate analysis. Functional status, measured by the Barthel score, was statistically significant in multivariate analyses, but did not satisfy the proportional hazards assumption. Pressure skin ulcers are also included in the final prediction models, and while this has not been reported in trajectory analyses, it has been reported in mortality models [15]. With regards to heart failure, it has been reported that frail patients are sometimes labelled with hearth failure diagnoses as a sign of their general state of decline, since standard classifications fail to recognise this patient condition [20]. Our results support this statement. Clinical severity and age are also reported as a main driver of time to death estimates.

This study has some limitations. First, the set of patient characteristics considered in this study may not be the same in other settings. For instance, other services may consider patients with different non-cancer diagnoses or that suffer less than three chronic illnesses. In addition, other services may use different variables to record clinical status. This may impact the generalizability of our study results. Furthermore, our analyses are based on baseline patient characteristics at programme enrolment, and only patient survival is considered in follow-up. Changes in functional, clinical status and other characteristics are not considered across time, as it usually done in trajectory analyses. Therefore, we do not know if or how patients transition from one illness trajectory to another. This remains as an area of future research.

Patient follow-up was stopped in January 2020 for the purposes of this study, due to the Covid-19 pandemic [25]. The mortality rates on patients included in this study due to Covid-19 diagnosis in the year 2020 would have biased survival estimates. Most patients included in this study were elderly and may have died due to Covid-19 instead of the underlying chronic illnesses they suffer. Comparability with data from previous years would have been at risk and so a decision was made to stop follow-up before the pandemic started and provide survival extrapolations instead of observing the complete curves.

An area of future research would be to explore healthcare utilisation and costs in each of the sub-groups included in this study. Identifying which services are most demanded among each group of patients may help in the organisation and planning of integrated health services.

CONCLUSION

This study complements the main end of life illness trajectories literature by reporting time to death estimates. The lack of survival predictions is a current obstacle for admitting non-oncologic patients into palliative care, for which we provide empirical evidence.

REFERENCES

1. Boyd, K. and S.A. Murray, Recognising and managing key transitions in end of life care. Bmj, 2010. 341: p. c4863.
2. Lynn, J. and D.M. Adamson, Living Well at the End of Life: Adapting Health Care to Serious Chronic Illness in Old Age. 2003: RAND Corporation.
3. Hajat, C. and E. Stein, The global burden of multiple chronic conditions: A narrative review. Preventive medicine reports, 2018. 12: p. 284-293.
4. Sambamoorthi, U., X. Tan, and A. Deb, Multiple chronic conditions and healthcare costs among adults. Expert review of pharmacoeconomics & outcomes research, 2015. 15(5): p. 823-832.
5. Ham, C., The ten characteristics of the high-performing chronic care system. Health Econ Policy Law, 2010. 5(Pt 1): p. 71-90.
6. Aira Ferrer, P., et al., Características de los pacientes con enfermedad crónica avanzada incluidos en un programa de cuidados paliativos domiciliario, in Medicina Paliativa. 2017. p. 179-187.
7. Fernández Moyano, A., et al., Care models for polypathological patients. Rev Clin Esp, 2017. 217(6): p. 351-358.
8. Amblàs-Novellas, J., et al., Identifying patients with advanced chronic conditions for a progressive palliative care approach: a cross-sectional study of prognostic indicators related to end-of-life trajectories. BMJ Open, 2016. 6(9): p. e012340.
9. Lunney, J.R., J. Lynn, and C. Hogan, Profiles of older medicare decedents. J Am Geriatr Soc, 2002. 50(6): p. 1108-12.
10. Cohen-Mansfield, J., M. Skornick-Bouchbinder, and S. Brill, Trajectories of End of Life: A Systematic Review. J Gerontol B Psychol Sci Soc Sci, 2018. 73(4): p. 564-572.
11. Ballentine, J.M., The Five Trajectories: Supporting Patients During Serious Illness. 2018, CSU Institute for Palliative Care.
12. Murray, S.A., et al., Illness trajectories and palliative care. Bmj, 2005. 330(7498): p. 1007-11.
13. Lynn, J., Serving Patients Who May Die Soon and Their FamiliesThe Role of Hospice and Other Services. JAMA, 2001. 285(7): p. 925-932.
14. Teno, J.M., et al., Dying trajectory in the last year of life: does cancer trajectory fit other diseases? J Palliat Med, 2001. 4(4): p. 457-64.
15. Bretos-Azcona, P.E., et al., Multisystem chronic illness prognostication in non-oncologic integrated care. BMJ Support Palliat Care, 2020.
16. Karon, I.H.B.A.H.I.D.E.G.Q.L.L.P.Z.F.S.L. Functional Trajectories At The End Of Life For Individuals With Dementia. 2020. January 2020.
17. Ibarrola Guillén C, Á.S., Udobro I I. . Estrategia Navarra de Atención Integrada a Pacientes Crónicos Y Pluripatológicos 2016; Available from: http://www.navarra.es/home_es/Temas/Portal+de+la+Salud/Ciudadania/Nuevo+Modelo+asistencial/Estrategia+navarra+de+atencion+integrada+a+pacientes+cronicos+y+pluripatologicos/.
18. Monterde, D., E. Vela, and M. Clèries, [Adjusted morbidity groups: A new multiple morbidity measurement of use in Primary Care]. Aten Primaria, 2016. 48(10): p. 674-682.
19. Bretos-Azcona, P.E., E. Sánchez-Iriso, and J.M. Cabasés Hita, Tailoring integrated care services for high-risk patients with multiple chronic conditions: a

risk stratification approach using cluster analysis. BMC Health Serv Res, 2020. 20(1): p. 806.
20. Lunney, J.R., et al., Patterns of functional decline at the end of life. Jama, 2003. 289(18): p. 2387-92.
21. Coventry, P.A., et al., Prediction of appropriate timing of palliative care for older adults with non-malignant life-threatening disease: a systematic review. Age Ageing, 2005. 34(3): p. 218-27.
22. Sercu, M., et al., Rethinking End-of-Life Care and Palliative Care: Learning From the Illness Trajectories and Lived Experiences of Terminally Ill Patients and Their Family Carers. Qual Health Res, 2018. 28(14): p. 2220-2238.
23. Grbich, C., et al., Identification of patients with noncancer diseases for palliative care services. Palliat Support Care, 2005. 3(1): p. 5-14.
24. Gill, T.M., et al., Trajectories of disability in the last year of life. N Engl J Med, 2010. 362(13): p. 1173-80.
25. Khan, S., et al., Emergence of a Novel Coronavirus, Severe Acute Respiratory Syndrome Coronavirus 2: Biology and Therapeutic Options. J Clin Microbiol, 2020. 58(5).

GENERAL THESIS CONCLUSIONS & FUTURE RESEARCH

GENERAL CONCLUSIONS

This thesis demonstrates, as a result of a risk stratification process, that the high-risk MCC patient population included in case management programmes is heterogeneous. Three distinct patient subpopulations within the top 5% risk stratum of the Kaiser pyramid are identified; each of them with their particular characteristics and needs. That is to say, we add additional layers to the Kaiser pyramid. These are *terminal illness*, *organ failure*, and *frail* patients, in ascending level of risk.

Therefore, we can say that "one size does not fit all", and that integrated care should be adapted to each of the resulting subgroups. While case management fits the needs of *terminal illness* patients, *organ failure* patients might need home-based care and *frail* patients would need palliative care.

From the risk score prediction model presented in Chapter 1, and the stratification from Chapter 2, several conclusions can be brought. First, it is the general severity of the patient what determines the level of risk of each individual patient and the risk strata to which they belong, not the number of simultaneous illnesses that a patient suffers. Two patients with different numbers of illnesses may be equally deserving, or equivalently, have the same level of risk.

In general, patients have greater levels of risk as they age, and decline in their functional, nutritional and renal status. In addition, pressure ulcers are associated with higher risk scores, as well as clinical severity and high nursing needs as measured by healthcare professionals. No particular diagnoses affected risk scores significantly, although it would be needed to investigate the effect of dementia in further analyses.

An issue worth mentioning is the development of the *nomogram* presented in Chapter 1, which allows an easy risk score calculation in clinical practice. This risk score calculation tool does not have problematical data requirements, as all information needed can be collected in question of minutes.

One of the novelties presented in this thesis is that Artificial Intelligence (AI) methods were used to stratify the population into different subgroups, based on clinical variables only. In particular, machine learning algorithms were used. Clinical variables were used to estimate risk scores, and this risk score was then used to determine how many patient subpopulations exist and the limits between them.

This is the very first time that such procedure has been done. In the past, AI methods were used, yet always using cost-based risk scores to stratify the population. Chapters 1 and 2 demonstrate that cost-based stratification tools, and GMA in particular, are poor risk adjusters when mortality or other clinical outcomes are considered. The GMA score was not a significant variable when predicting death, and did not vary significantly between patient subgroups. However, mortality rates and risk scores did vary significantly between clusters.

Cost-based stratification tools have a lack of alignment between their purpose, which stratify patient populations based on resource utilisation, and the purpose of this study, that is, to stratify patients according to their clinical needs and then adapt care strategies accordingly. Stratifying populations based on clinical risk scores ensure that clinical needs and not cost or financial needs are addressed. Therefore, our clinical-based risk stratification process is more advantageous in terms of the objectives of this thesis.

One of the main barriers towards palliative care programme's implementation for non-oncologic chronic patients is that determining the moment in which a patient starts this end-of-life phase is complicated, due to entry-re-entry episodes of illness. In this situation, all episodes of illness can potentially result in death, yet patients typically suffer several of such episodes and then recover. Healthcare professionals have argued that it is virtually impossible to know when the final phase of life will start as any of these entry-re-entry episodes could be the very last one. End-of-life phase in non-oncologic MCC patients is said to be non-predictable. The consequence is that palliative care is usually delayed, and most of the times not started.

This thesis demonstrates in Chapter 3 that survival, and time to death estimates, are indeed predictable in non-oncologic MCC patients. This has important implications as the results presented in this document help to overcome the obstacles in palliative care implementation. Non-oncologic MCC patients should no longer be excluded from this type of services, as it is now possible to estimate the time to death for each individual patient in our population.

We also find prognostic variables that predict time to death. The recognition of predictors constitutes the starting process of early identification of palliative patients' needs, and guides transitions towards end-of life care. In addition, it helps to develop

objective criteria for a correct and timely identification of those patients in whom the terminal phase of their life will happen in the near future.

All in all, this thesis seeks to improve health outcomes and the care provided to high-risk CCM patients currently included in the case management program. The risk stratification achieves this purpose, in addition to seeking care that adapts better to the needs of each type of patient. Specifically, special emphasis is placed on the need to improve the functional capacity and quality of life of patients, while adapting care to the individual level of dependency and autonomy. Similarly, organizational improvements are suggested in integrated care, supported by the evidence presented here.

FUTURE RESEARCH

Among the future lines of research that should be followed, it is necessary to investigate whether the results presented in this thesis are still valid in populations of chronic patients with fewer than three simultaneous diseases, or with different diagnoses.

In this thesis it has been determined that there are three subgroups of patients and their estimated time to death. However, it has not been possible to determine whether or when patients move from one group to another. This needs to be investigated to optimize the design of new care programs.

It is also necessary to explore the patterns that non-oncological CCM patients follow with respect to the use of resources and costs. In the same way that we have determined the clinical needs for each type of high-risk patient, it is equally important to know what type of resources are used most intensively in each subgroup. This will further help the design of new integrated care programs.

In the event that new integrated care programs are implemented for each subgroup, it will be necessary to determine the improvement in efficiency. In other words, perform cost-effectiveness studies.

Finally, it would be interesting to investigate the effect of external shocks on the non-cancer high-risk MCC population, such as COVID-19, a heat wave or possible seasonal effects.

SUPPLEMENTARY MATERIALS

SUPPLEMENTARY MATERIAL 1: MISSING DATA & MULTIPLE IMPUTATION

Missing data

Missing data occur very frequently in almost all health-related research that involves individual patient data. The inadequate management of this can produce biased estimates of parameters both in descriptive statistics, and in other statistical estimations such as regression analyses.

The alternative to data imputation, which is simply to delete all observations with missing values, implies losing all the information registered in non-missing variables for these observations. This results in situations where the remaining complete observations are not representative of the population of interest and biased results are generated as a consequence. It is important to use all the information that is available so as to not discard any information of interest.

Missing data can be classified into three different types of data depending on the patterns of missing data:

1) Missing completely at random (MCAR)

   Data are MCAR if the probability that data are missing does not depend on observed or unobserved data. That is to say, the fact that data are missing are unrelated to any characteristics of the study participants or design. In this circumstance, discarding observations does not lead to biased estimates, but maybe less powerful results.

2) Missing at random (MAR)

   In this case, the probability that data are missing may depend on observed data. In other words, the amount of information collected depends on a 3rd variable that is included in the dataset. For example, in our case data are MAR if more severe individuals are more likely to have their clinical variables recorded (and severity is included in the analysis). As individuals worsen their status, more tests and doctor visits will happen. As a consequence, more information will be collected. When missing data are MAR, deleting observations may lead to biased results.

3) Missing not at random (MNAR)

If data are MNAR, the probability of data being missing depends on unobserved data, conditional on the observed data. The availability of data depends on variables that are no longer observed once the participants are excluded from analyses.

Multiple imputation techniques handle missing data and are increasingly popular. The key idea of multiple imputation is that it estimates a set of possible values for the missing data that we desire to complete. A MAR assumption is needed to perform multiple imputation. In our case, given that the availability of information is assumed to be greater as patient severity increases, and severity is indeed included in the dataset, it is safe to assume that data are MAR.

The process of generating multiple imputation data goes as follows:

First, *m* multiple imputed data sets are created, where missing data are replaced by simulated values drawn from the distribution of missing data conditional on the observed data. A key aspect of MICE is to decide how many imputations should be estimated. The literature often suggests using values for m=5, as this corresponds to an efficiency of 95% if there are less than 50% of missing data.

For a variable *z* that contains missing values, a regression model that regresses *z* on the set of variables that are already complete is performed. However several variables with missing values should be imputed, as it is our case. In this situation, multiple imputation by chained equations (MICE) is a useful method.

MICE produces imputed values for each variable with missing values. All missing values are first filled in using random sampling with replacement from the observed values of each variable to be imputed. Then, the variable with the lowest percentage of missing values is regressed on all other variables, restricting the regression to the observations with observed values for that particular variable. The randomly-imputed values are then replaced by the new estimated values. The process continues with the next variable until all variables have been updated in a cycle. This cycle is repeated many times in an iterative way (as many as necessary) to produce a single set of imputed values until results stabilise (generally the mean and standard deviation of the imputed variable).

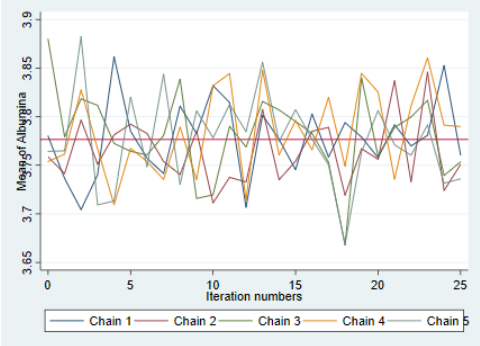
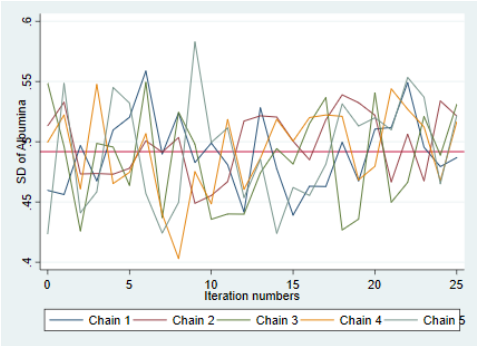
MICE takes into account the different nature of each variable (continuous, binary, categorical), and allows the selection of the particular regression method that is applied to each variable. In this way, each variable is imputed using its own model.

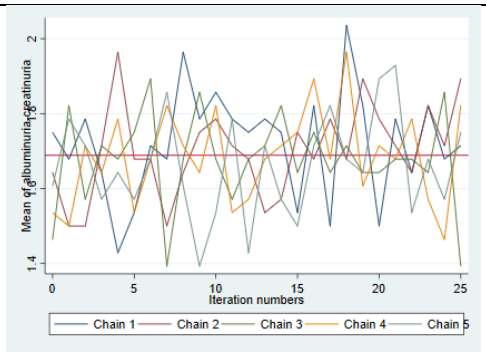
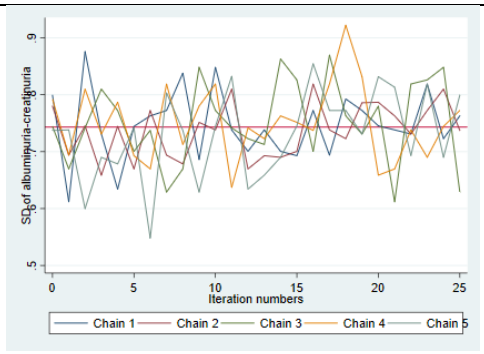
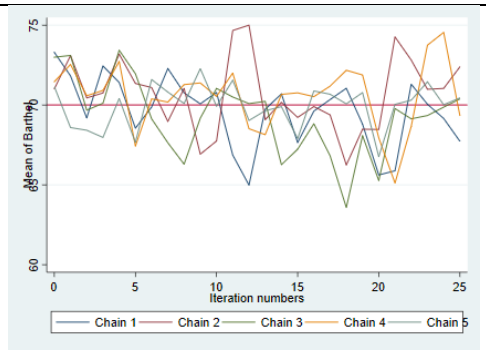
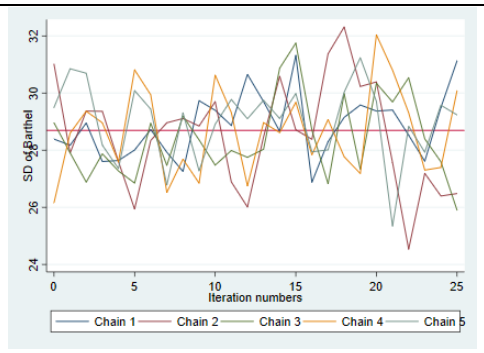
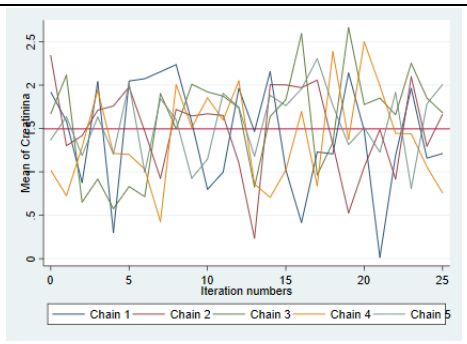
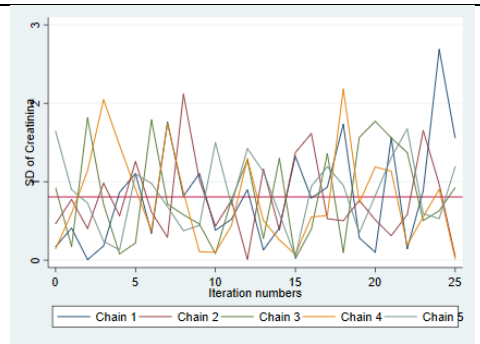
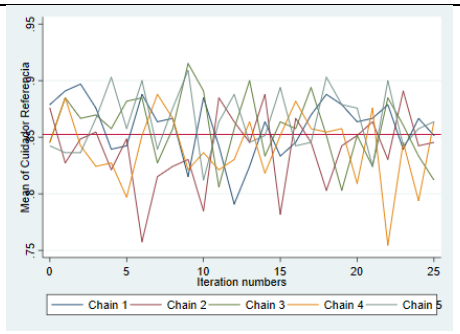
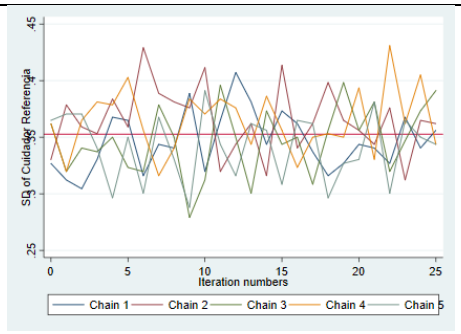
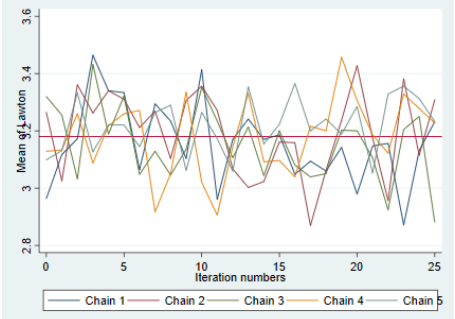
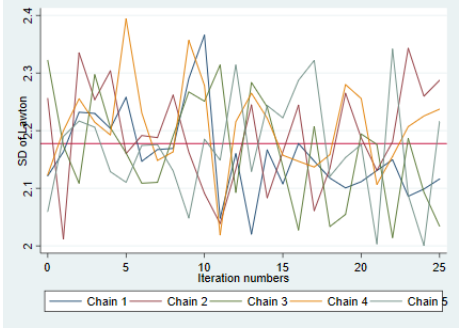
Since MICE is an iterative procedure in which several imputation cycles are performed, it is important to assess whether convergence is achieved around the mean and standard deviation of the imputed variables. The number of cycles can be chosen, and convergence can be assessed visually, as shown below for the particular case of the data set used in this thesis.

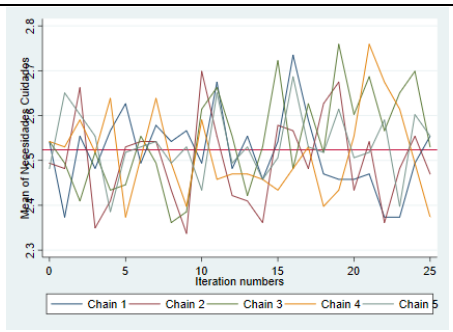
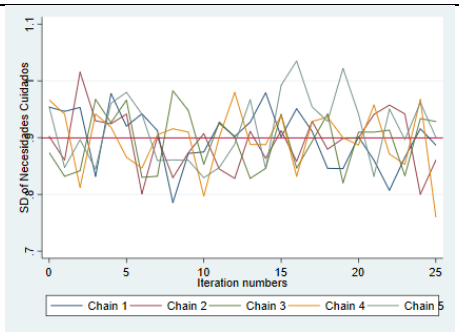
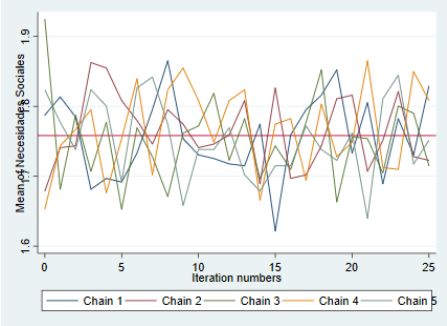
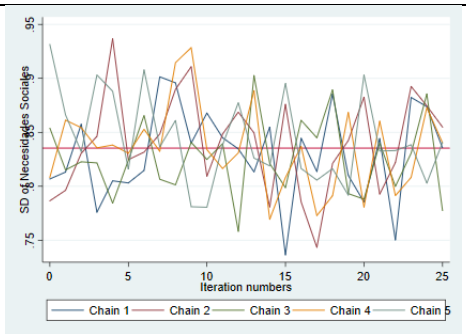
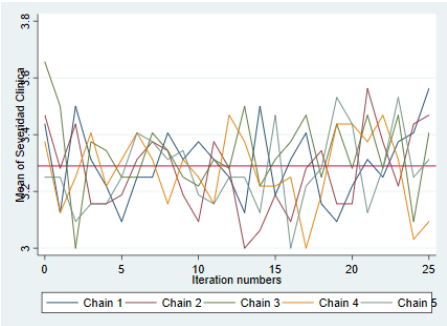
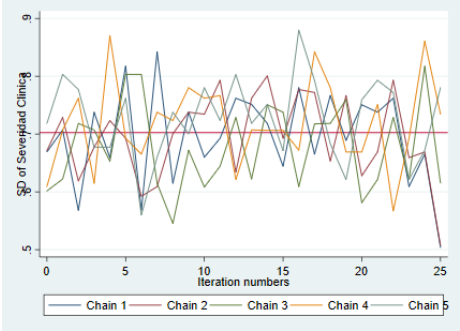
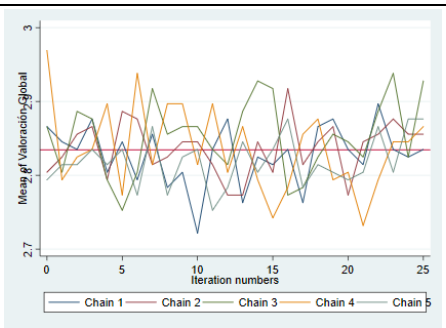
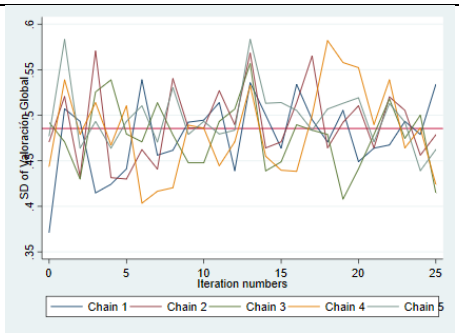
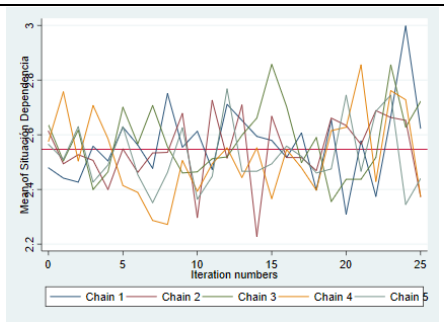
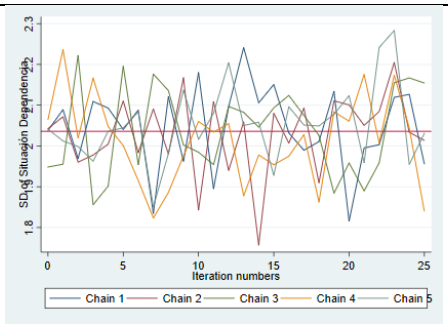
The second step is to use the imputed data in the analyses. The analyses (i.e. a regression) are repeated as many times as *m* imputed data sets exist. The analyses are repeated identically for each imputed data set. Finally, all results are pooled into a final result, such as regression coefficients or confidence intervals.

Assessing MICE convergence

In the graphs presented below, we can observe how all variables that contain missing values achieved convergence in their mean and standard deviation values across all 5 imputations that were performed. 25 iterations are shown for each variable, and in all cases convergence was achieved at that point. We can therefore consider that all variables were correctly imputed, and that we can use these values in our analyses.

Variable	Mean	Standard Deviation
Albumin		

Albumin/ creatinine index		
Barthel scale		
Creatinine		
Informal carer		
Lawton & Brody scale		

Nursing needs		
Social needs		
Clinical severity		
Global severity		
Degree of dependency		

SUPPLEMENTARY MATERIAL 2: LOGISTIC REGRESSION – PERFORMANCE MEASURES AND VALIDATION

The logistic regression model is the most widely used statistical technique when facing binary outcomes, as it is the case in Chapter 1. The binary outcome Y is linked to a linear combination of a set of predictors and regression coefficients β. The logistic link function is used to restrict predictions to the interval [0,1]. The model is stated in terms of the probability that Prob.(Y=1), and the β coefficients are usually estimated by maximum likelihood.

$$\text{Prob}(Y_i = 1) = F(\beta_0 + \beta_1 X_i) = \frac{e^{\beta_0 + \beta_1 X_i}}{1 + e^{\beta_0 + \beta_1 X_i}} = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_i)}}$$

The interpretation of the β coefficients is as for any regression model. The coefficient indicates the effect of a one-unit increase in the predictor of interest, keeping the other predictors in the model constant. It is possible to relate the coefficients in the model to the odds of an event occurring, which is usually the measure reported (as it is done in Chapter 1). Odds denote number of successes for every failure.

$$\text{Odds}(Y_i = 1) = \frac{\text{Prob}(Y_i = 1)}{1 - \text{Prob}(Y_i = 1)} = \in (0, \infty)$$

The exponent of the regression coefficients indicate the odds ratios. Therefore the coefficients are just the natural logarithm of the odds ratios. Below are reported the final models from Chapter 1 in the form of coefficients rather than Odds ratios:

Multivariate Logistic Regression to 1-Year Mortality.

	Model 1		Model 2	
Variable	**Coefficients**	**p-value**	**Coefficients**	**p-value**
Barthel Scale	-0.0208	0.000	-	-
Degree of Dependency II: Severe	-	-	0.8515	0.016
Degree of Dependency III: Great Dependency	-	-	0.9714	0.001
Creatinine (mg/dL)	0.2499	0.032	0.2546	0.027
Presence of Pressure Ulcers	-	-	0.4524	0.029
Severe Global Status	1.2734	0.000	1.3425	0.000
Very Severe Global Status	3.0301	0.000	3.0840	0.000
Constant	-1.1340	0.007	-2.8658	0.000
Area Under the Curve	0.751		0.737	

Dependent Variable: 1-Year Mortality; CI: Confidence Interval

Model discrimination ability

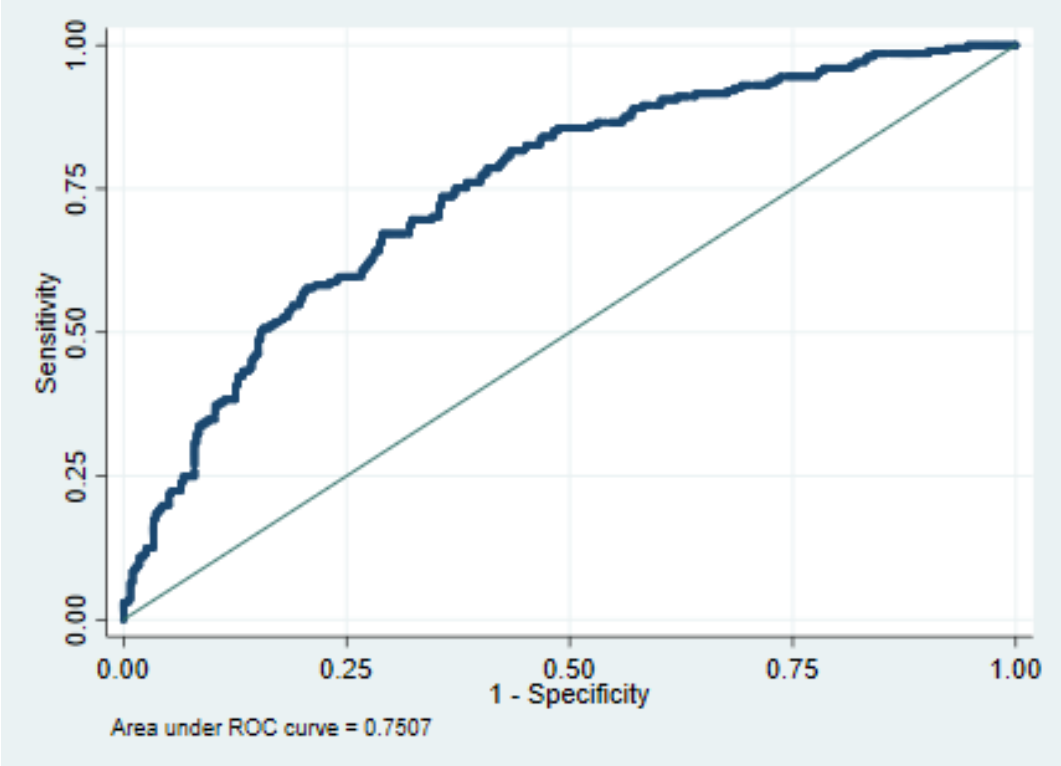
The ability of a model to discriminate between those with and those without the outcome is measured using the area under the receiver operating characteristic (ROC) curve (AUC).

Usually, to classify a patient as positive (having the outcome) or negative (not having the outcome), it is needed to apply a cut-off value to the predicted model probability of suffering the outcome. If the prediction is superior to the defined cut-off value, for example 50%, it will be classified as positive, and it the prediction is lower, it will be negative.

The ROC graphs the sensitivity (true positive rate) against 1 – specificity (false-positive rate) for all possible probability cut-offs from 0% to 100% for the prediction of outcome of interest in our model. The AUC quantifies the discrimination ability of a particular model in a scale 0-1. An AUC=0.50 will indicate a non-informative model, and an AUC=1 will indicate a model that discriminates perfectly.

Below you will find the ROC for model 1 of Chapter 1, which demonstrated to be the model with a better discrimination ability (AUC=0.751).

Receiver Operating Characteristic (ROC) of model 1, Chapter 1

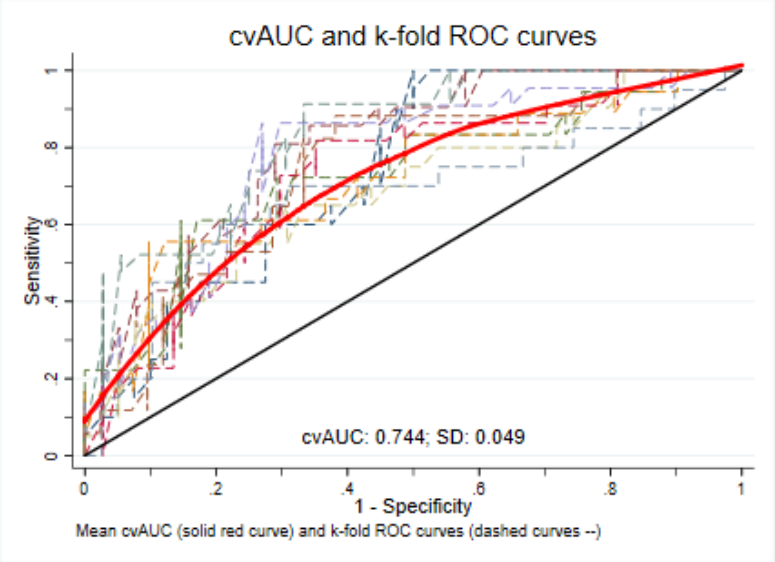


Validation

Cross-validation (K-fold) techniques were used to validate the model internally. These test the model on a random part that is left out of the sample. The model is run on the part of the data that remains, and tested on the random part of the data that was left out for that purpose. The purpose is to check whether the discriminative ability of the model, measured using the ROC/AUC methods still hold even if leaving random observations out. This process is repeated for consecutive fractions of the sample, let's say 10 fractions each containing 10% of the data each. The model will be tested in 1 of the 10 random fractions, and the process is repeated in each of the folds. The global performance is the average of all AUC values.

In this thesis, we perform a 10-fold cross-validation process on model 1 of Chapter 1, as shown below, which resulted on an AUC=0.744.

K-fold validation for model 1, Chapter 1



We also validated the model using bootstrapping methods, which samples from the underlying population, with replacement. That is to say, samples are drawn from the complete dataset to validate the model using the AUC, and then put back into the full dataset. Then, at each bootstrap, the same observations may be drawn again. This is the key difference in comparison to cross-validation, where all random folds are independent between them, and all observations are included into the validation process only once. Here, the same observation may be randomly drawn several times, or none as it is put back in the original dataset after validating results on that particular bootstrap. We use 1000 bootstraps, which yield an AUC= 0.751.

SUPPLEMENTARY MATERIAL 3: CLUSTER ANALYSES – HIERARCHICAL ALGORITHMS AND OPTIMAL NUMBER OF CLUSTERS

Chapter 2 uses cluster analysis techniques to identify groups in data. These processes are also known in the artificial intelligence literature as unsupervised machine learning, or simply segmentation. Usually, cluster analyses are performed using different types of algorithms, for which additional details are provided below to complete the information provided in the main thesis text.

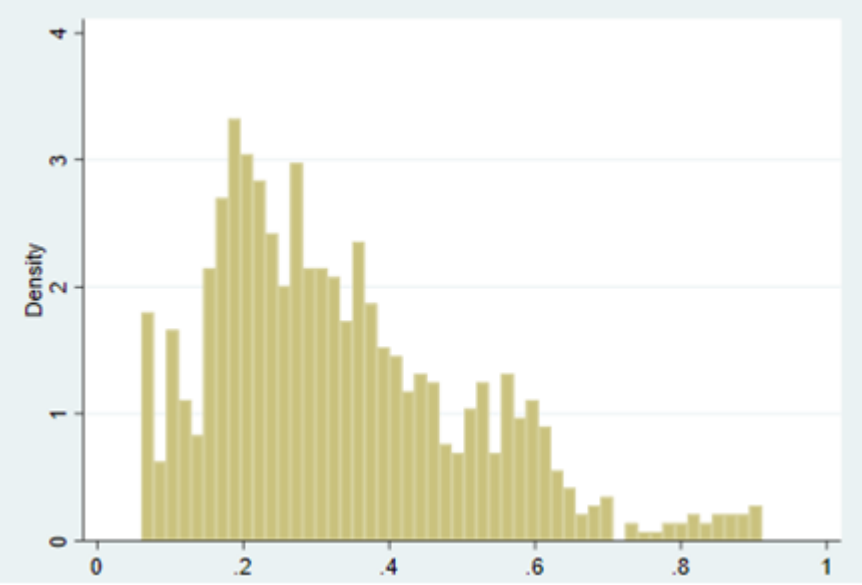
Detecting clusters graphically

Before starting to apply algorithms directly to the data, making a visual inspection of the data may be useful to suggest that the data does indeed contain different types of observations and consequently that some additional clustering methods may be useful in uncovering different population subgroups.

If univariate data are used in the clustering process, as it is our case where a risk score is the only variable used in the segmentation, a visual inspection will be useful to detect possible data patterns. A unimodal histogram will indicate a homogeneous population without subgroups, whereas multimodal histograms will indicate that the population is heterogeneous. In this case, every 'bump' will indicate a possible subgroup.

As disclosed in the histogram below, which is representing the estimated risk scores in our population of high-risk MCC patients, we can distinguish the existence of three possible clusters. The first one up to a probability of ~ 0.5, then up to ~ 0.7 and the rest.

Histogram of the predicted 1-year mortality rates



Hierarchical clustering

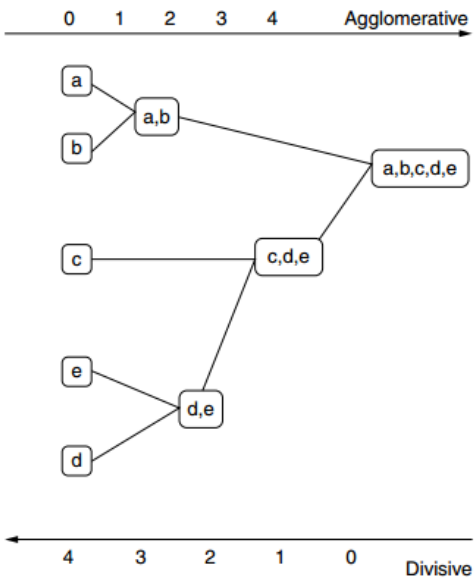
Since we already have the intuition of the existence of patient subgroups in the population, it is now worth proceeding with proper clustering methods. One of the most common type of algorithms used are hierarchical algorithms.

Hierarchical classifications do not produce a data segmentation in a single step. Instead, the segmentation is produced in a series of steps that can be either divisive or agglomerative.

Agglomerative procedures consist of a series of successive fusions of the individual observations into bigger groups. In this way, the first step consists of *n* single-member clusters, and the last consists of a single group that contains all observations. In contrast, divisive methods separate a groups that contains all individuals into finer groups until all observations have been separated into *n* single-member clusters.

Hierarchical clustering processes can be represented in a type of diagram known as a dendrogram, which illustrates the mergers or divisions made at each stage of the analysis depending if an agglomerative or divisive method was used.

Hierarchical clustering methods: Dendrogram



Hierarchical methods rely on the distance between observations to do the mergers or divisions. For example, agglomerative methods merge those observations or subgroups that are closer between them according to a specific criterion.

In this thesis, Ward's linkage criterion is used to determine which clusters should be merged. Ward's linkage determines that the fusion of two clusters should be based on the minimisation of the increase of the size of an error sum-of-squares measure E, which is assessed at each fusion stage. E is expressed as:

$$\mathrm{E} = \sum_{\mathrm{m}=1}^{\mathrm{g}} \mathrm{E_m}$$

Where

$$\mathrm{E}_m = \sum_{l=1}^{n_m} \sum_{k=1}^{p_k} (x_{ml,k} - \bar{x}_{m,k})^{\,2}$$

Where *m* stands for the *m*th cluster and *k* for the *k*th variable for the *l*th object.

The increase of the size of E is proportional to the squared Euclidean distance between the centroids of the newly merged clusters. The Euclidean distance is expressed as:

$$d_{ij} = \left[\sum_{k=0}^{p} (x_{ik} - x_{jk})^2 \right]^{1/2}$$

Optimal number of clusters

When using agglomerative hierarchical clustering methods, the process will always end up with a single group containing all observations. It is then needed to stop, or 'cut' the dendrogram at a point in the merging process deemed as optimal. This will provide the final number of subgroups that results from the clustering process.

There are several methods to find out the optimal number of clusters or subgroups. For instance, the 'Duda and Hart' stopping rule that requires hierarchical clustering information, as it is the case in Chapter 2. The rule uses a criterion which is the Je(2)/Je(1) index and its associated pseudo-T -squared value. Je(1) is the sum of squared errors within the resulting subgroups of the clustering process. Je(2) is the sum of squared errors in the resulting subgroups. A large Je(2)/Je(1) index value and a small pseudo-T -squared value indicate an optimal number of clusters. The relationship between the Je(2)/Je(1) index and the pseudo-T -squared value is given by:

$$\frac{1}{Je(2)/Je(1)} = 1 + \frac{T^2}{N_1 + N_2 - 2}$$

Duda and Hart stopping rule results

Number of clusters	Je(2)/Je(1)	pseudo T-squared
1	0.3804	1437.96
2	0.3161	1538.48
3	**0.2741**	**499.03**
4	0.2754	1126.35
5	0.2779	730.08
6	0.2736	815.24
7	0.3489	268.71
8	0.2476	537.88
9	0.1882	513.15
10	0.4602	97.38

As observed in the table above, the optimal number of clusters is three, in line with the graphical solution also observed at the risk score histogram. At three clusters, a large Je(2)/Je(1) index value and a small pseudo-T -squared value are first observed.

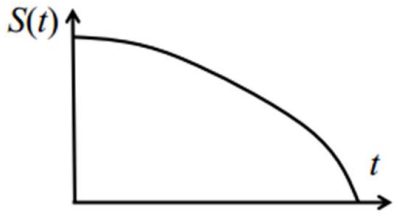
SUPPLEMENTARY MATERIAL 4: SURVIVAL ANALYSES – PROPORTIONAL HAZARDS ASSUMPTION AND HAZARD RATES

Survival function and the hazard rate

The survival function S(t) describes the probability of survival beyond time *t*, where S(t) can be defined as:

$$S(t) = Prob.(T > t) = 1 - \int_0^t f(s)ds$$

Survival function S(t)



The rate at which events (deaths) occur per unit of time can be expressed with the hazard rate h(t), which shows the instantaneous probability of an event occurring, or equivalently the probability of the event occurring in the next short time interval, conditional on non-occurrence before time *t*. h(t) can be expressed as:

$$h(t) = \lim_{\Delta t \to 0} \frac{pr(t \leq T \leq t + \Delta t | T > t)}{\Delta t} = \frac{-S'(t)}{S(t)}$$

Survival models, hazard rates and coefficients

The taxonomy of survival analyses distinguishes between two broad types of models if they incorporate information on individual characteristics as predictors. These are proportional hazards (PH) models, and accelerated failure time models. The key assumption for the proportional hazards type of model is that individual characteristics shift up or down the hazard rate, whereas accelerated failure time models speed up or down the time to event depending on said individual characteristics.

Therefore, PH models assume that there is a common baseline survival for all types of patients, or patient subgroups, and that this baseline survival shifts up or down depending on individual characteristics. For any explanatory characteristic, the

proportional difference in hazard rates between those with the characteristic and those without the characteristic is constant. In other words, not a function of survival time.

Let Z_i be a vector of characteristics for an individual *i*, and $h_0(t)$ the baseline hazard function, which is a function of *t* but not Z, and which summarises the pattern of survival common to all. Then the PH type of models assume that the hazard for the individual *i* is given by:

$$h(t, Z_i) = e^{\beta' Z_i} h_0(t) = \lambda_i h_0(t)$$

where

$$\lambda_i = e^{\beta' Z_i}$$

and

$$\beta' Z_i = \beta_0 + \beta_1 Z_1 + \beta_i Z_i$$

Ceteris paribus, the effect of a variable on the hazard rate can be expressed as:

$$HR(Z_2 = 1) = \frac{h(t, Z_2 = 1)}{h(t, Z_2 = 0)} = \frac{e^{(\beta_0 + \beta_1 Z_1 + \beta_2 1)} h_0(t)}{e^{(\beta_0 + \beta_1 Z_1 + \beta_2 0)} h_0(t)} = e^{\beta_2}$$

In the Chapter 3 main text, results are expressed as hazard ratios, yet it is possible to report results as coefficients. This is particularly interesting for the results that were reported, given that the constant term showed an hazard ratio of 0, yet it was not exactly 0. In this case, the rounding did not show the (very small) decimals. However, they can be seen if coefficients are reported instead.

In order to transform HR into coefficients, the following transformation is used:

$$e^{\beta_i} = \mathrm{HR}(\Delta Z_i = 1)$$

Alternatively:

$$\beta_i = \log(\mathrm{HR}(\Delta Z_i = 1))$$

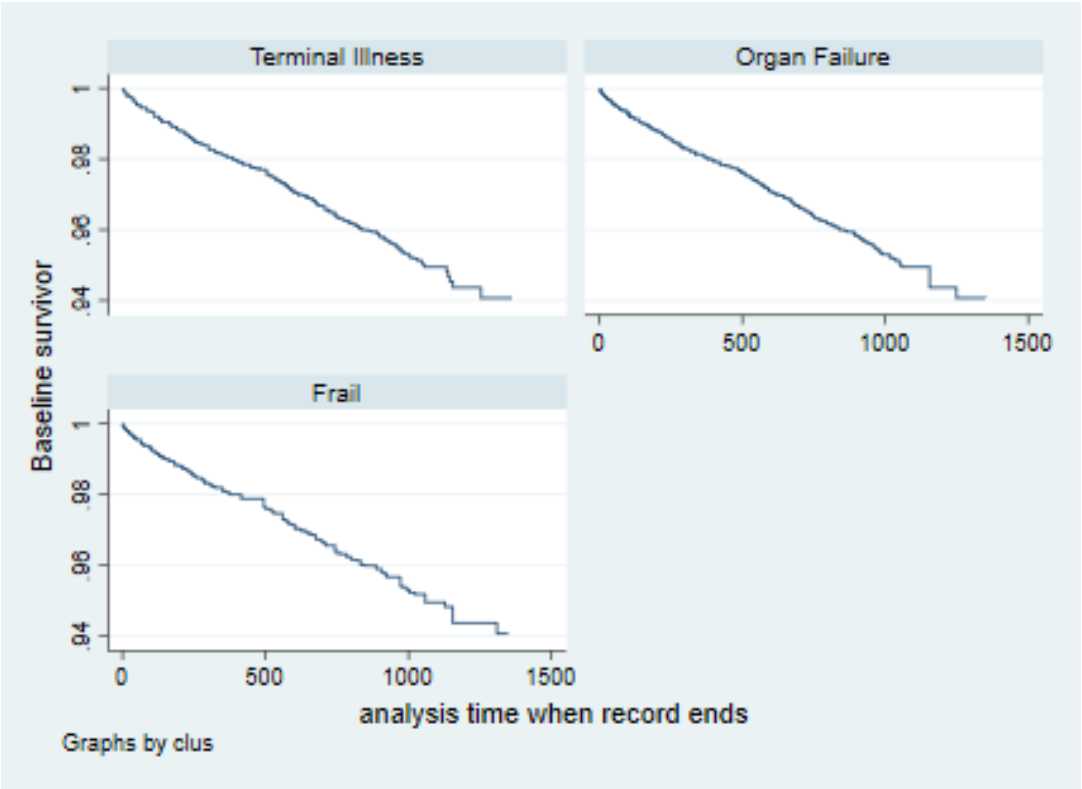
Below are the results from the Weibull regression presented in Chapter 3, in the form of coefficients rather than HR.

Weibull regression	Hazard rate	p-value	95% Conf. Interval
Age	0.4182	0.000	0.0286 - 0.0550
Pressure skin ulcers	0.3958	0.000	0.1977 – 0.5938
Heart failure	0.2551	0.018	0.0446 – 0.4657
Severe global severity	0.8153	0.000	0.4713 – 1.1591
Very severe global severity	1.8758	0.000	1.3944 – 2.3571
Constant	-10.3635	0.000	-11.6315 - -9.0954
ancillary parameter p	0.796		0.730 - 0.866
1/p	1.257		1.155 - 1.368
ancillary parameter p: Terminal illness	1.032		0.879 - 1.212
ancillary parameter p: organ failure	0.755		0 .662 - 0.860
ancillary parameter p: frail	0.754		0.648 - 0.877

Testing the proportional hazards assumption

It is important to test whether the PH assumption holds to build either a PH model, or an accelerated failure time model. In the Chapter 3 main text, it is explained that the Schoenfeld residuals test was used.

However, it is also possible to plot the baseline survival function h(0) for each of the subpopulations of interest. After a visual inspection, if they all look similar we can consider that they are proportional and that a PH model can be fit.



As we observe above, all three baseline survival functions for the three subgroups of interest are similar. Therefore, based on this evidence and the Schoenfeld residuals tests presented in Chapter 3 we can consider a good decision to fit a PH model. Several PH distributions were tested, with the Weibull distributions showing the best fit. The Weibull model also has an accelerated failure time form, yet the PH was the one used in this thesis.

SUPPLEMENTARY MATERIAL 5: STATA CODE

The most relevant STATA code used in this thesis is presented below:

1) Multiple Imputation steps

```
mi set wide

*prove that the data are not monotone

mi misstable patterns
mi misstable nested
mi misstable summarize

*register variables with missing values
mi register imputed Sit_Dependencia necesidades_sociales necesidades_cuidados severidad_clinica valoracion_global albumina creatinina cuidador_referencia barthel lawton albuminuria_creatinur microalbuminuria

mi describe

*Imputation step

mi impute chained
(ologit, include(Sit_Dependencia albumina creatinina cuidador_referencia barthel lawton albuminuria_creatinur microalbuminuria) noimputed) necesidades_sociales
(ologit, include(Sit_Dependencia albumina creatinina cuidador_referencia barthel lawton albuminuria_creatinur microalbuminuria) noimputed) necesidades_cuidados
(ologit, include(Sit_Dependencia albumina creatinina cuidador_referencia barthel lawton albuminuria_creatinur microalbuminuria) noimputed) severidad_clinica
(ologit, include(Sit_Dependencia albumina creatinina cuidador_referencia barthel lawton albuminuria_creatinur microalbuminuria severidad_clinica necesidades_cuidados necesidades_sociales) noimputed) valoracion_global
(ologit, include(Sit_Dependencia albumina creatinina cuidador_referencia barthel albuminuria_creatinur microalbuminuria severidad_clinica necesidades_cuidados necesidades_sociales valoracion_global) noimputed) lawton
(ologit, include(Sit_Dependencia albumina creatinina cuidador_referencia lawton albuminuria_creatinur microalbuminuria severidad_clinica necesidades_cuidados necesidades_sociales valoracion_global) noimputed) barthel
(ologit, include(albumina creatinina cuidador_referencia barthel lawton albuminuria_creatinur microalbuminuria severidad_clinica necesidades_cuidados necesidades_sociales valoracion_global) noimputed) Sit_Dependencia
(ologit, include(Sit_Dependencia albumina creatinina cuidador_referencia barthel lawton severidad_clinica necesidades_cuidados necesidades_sociales valoracion_global) noimputed) microalbuminuria
(ologit, include(Sit_Dependencia albumina creatinina cuidador_referencia barthel lawton severidad_clinica necesidades_cuidados necesidades_sociales valoracion_global) noimputed) albuminuria_creatinur
(logit, include(Sit_Dependencia albumina creatinina barthel lawton albuminuria_creatinur microalbuminuria severidad_clinica necesidades_cuidados necesidades_sociales valoracion_global) noimputed) cuidador_referencia
(regress) creatinina albumina
 = Pac_Fallecido Pac_Edad Sexo_Code GMA_weight comorbidities ulcera_cronica_piel, rseed(5858) add (10) dots

*checking convergence

use impstats, clear
summ

reshape wide *mean *sd, i(iter) j(m)

tsset iter

tsline creatinina_mean1 creatinina_mean2 creatinina_mean3 creatinina_mean4 creatinina_mean5 , ytitle(Mean of Creatinina) yline(1.49576) legend(rows(1) label(1 "Chain 1") label(2 "Chain 2") label(3 "Chain 3") label(4 "Chain 4") label(5 "Chain 5"))

*the process is repeated with all imputed variables' mean and standard deviation
```

2) Logistic regression

```
mi estimate, saving (miest, replace): logit fallece_1aÃ±o barthel creatinina val_glob_severo val_glob_cravanzado
 mi predict y_3 using miest
quietly mi xeq: generate Y_3= invlogit(y_3)
histogram Y_3
*obtengo AUC & gráfica
mi estimate, cmdok: eroctab fallece_1aÃ±o Y_3
roctab fallece_1aÃ±o Y_3, graph msize(tiny)
*Validation
cvauroc fallece_1aÃ±o AVG_barthel avg_creatinina AVG_val_glob_severo AVG_val_glob_cravanzado, kfold(10)
seed(123) fit graphlowess
rocreg fallece_1aÃ±o Y_3
```

3) Clustering process

```
cluster wardslinkage Y_3
cluster stop, rule (duda)
cluster dendrogram, cutnumber (100)
*El dendrograma sugiere 3 sub-grupos Y Duda Hart tambien

*I create 3 clusters
cluster generate clus_Y_3 = groups(3), name(_clus_1) ties(error)
tab clus_Y_3

*Non-hierarchical methods
* K-MEANS
cluster kmeans Y_3, k(3)
```

4) Survival analyses

```
gen time=0
replace time= (Fecha_Muerte-fecha_activacion_cubo_rojo) if Fallecimiento==1
replace time= (end_study-fecha_activacion_cubo_rojo) if Fallecimiento==0

codebook time
tab time if time<100
replace time=0.5 if time==0
summ time

*Tell Stata that our dataset consists of survival time data
*I've adapted the command to MI data

mi stset time, failure(Fallecimiento==1)
```

```
*genero variable S(t)
sts gen surv=s, by(clus)

*hago el grafico de S(t)
graph twoway scatter surv time, msize(tiny)
sepscatter surv time, separate(clus) msize (tiny) ylabel (0 (0.25) 1)

*miro la gente que hay en cada cluster, su supervivencia, su tiempo, cuantos muertos y censored
sts list, by(clus)

*MEAN survival time (restricted + extended)
stci, by(clus) rmean
stci, by(clus) emean

*log-rank test
sts test clus, logrank

*MEDIAN survival time
stci, by(clus) median
stsum, by(clus)
*graph
sts graph, risktable xlabel (0(100)1400, alternate)
*(1) Modelling
stepwise, pr(0.05): stcox Pac_Edad AVG_barthel ulcera_cronica_piel AVG_albumina avg_creatinina fpsi2
          insuficiencia_cardiaca demencia cirrosis AVG_Val_G_Sev AVG_Val_G_Very_Sev AVG_severidad_mod
          AVG_severidad_sev AVG_severidad_very_sev AVG_nec_leve AVG_nec_mod AVG_nec_sev

************************************
* PROPORTIONAL HAZARDS ASSUMPTION *
************************************

*checking the proportional hazards assumption

*Working with the baseline hazard
predict base_surv_cox, basesurv
line base_surv_cox _t, sort c(J)
graph twoway scatter base_surv_cox _t, msize(tiny) sort c(J)
graph twoway line base_surv_cox _t, sort c(J) by(clus)
*All groups have the same baseline hazard pattern, this is good

*COX&KM
stcoxkm, by(clus) ylabels(0 0.25 0.5 0.75 1)
stcoxkm, by(clus) separate ylabels(0 0.25 0.5 0.75 1)
```

```
stcoxkm, by(clus) ylabels(0 0.25 0.5 0.75 1)legend(label(1 "Terminal ilness") label(2 "Organ failure") label(3 "Frail")
        label(4 "Predicted: Terminal illness") label(5 "Predicted: Organ failure") label(6 "Predicted: Frail"))
* plots Kaplan-Meier observed survival curves and compares them with the Cox predicted curves for the same
        variable. The closer the
*observed values are to the predicted, the less likely it is that the proportional hazards assumption has been
        violated

*regression:
stcox Pac_Edad AVG_barthel ulcera_cronica_piel insuficiencia_cardiaca AVG_Val_G_Sev AVG_Val_G_Very

*PH Test
estat phtest, detail
*problem; does not hold in the table

*weibull
mi estimate, saving (miest, replace): streg Pac_Edad ulcera_cronica_piel insuficiencia_cardiaca val_glob_severo
        val_glob_cravanzado, d(weibull) nolog
mi estimate, hr
streg Pac_Edad ulcera_cronica_piel insuficiencia_cardiaca AVG_Val_G_Sev AVG_Val_G_Very, d(weibull) nolog
estat ic
```

LAST PAGE