

Florence DUCHÊNE
INPG – DEA SIPT 2001



Stage encadré par Mr Vincent RIALLE
1^{er} mars – 30 juin 2001

INPG

46, avenue Félix Viallet
38031 Grenoble Cedex 1
Tél. 04 76 57 45 00 - Fax. 04 76 57 45 92

Rapport de stage

Exploitation de données multicateurs dans un système de télésurveillance médicale de patients à domicile

**Application au projet HIS :
Habitat Intelligent pour la Santé**



**Laboratoire TIMC / IMAG – UMR CNRS 5525
Equipe AFIRM**

Université Joseph Fourier
Faculté de Médecine de Grenoble
38706 La Tronche Cedex

Remerciements

Je tiens à remercier Monsieur Jacques DEMONGEOT qui m'a accueillie au sein de son laboratoire dans le cadre de mon stage de DEA, ainsi que Monsieur Norbert NOURY qui m'a intégrée dans son équipe de recherche et s'est intéressé de près à mon travail.

Je suis reconnaissante envers Monsieur Vincent RIALLE, mon responsable de stage, pour avoir suivi mon travail durant ces quelques mois.

Je remercie aussi tous les membres de l'équipe pour leur accueil et leur soutien, et en particulier Monsieur Lionel BAJOLLE pour ses compétences médicales qui m'ont été indispensables et Monsieur Gilles VIRONE pour sa connaissance du projet qu'il m'a transmise.

Un grand merci aussi à Monsieur Christophe BERENGUER, maître de conférences à l'Université de Technologie de Troyes, pour le temps qu'il m'a consacré et l'intérêt qu'il a porté à mon sujet.

Enfin un gros clin d'œil à tout mon entourage et en particulier à mes parents pour leur disponibilité et leur soutien tout au long de cette année charnière pour la réorientation de ma carrière professionnelle.

Résumé

La télésurveillance médicale représente un enjeu considérable face au vieillissement de la population et à la pénurie de médecins pressentie pour les années à venir. Ce document présente une étude autour du concept d'habitat intelligent pour la santé. Il s'agit d'équiper un appartement de divers types de capteurs et d'analyser en temps réel les signaux reçus en s'appuyant sur d'autres informations de type clinique par exemple pour détecter les situations inhabituelles et urgentes d'une personne à domicile. Cet habitat est intégré à un réseau médical qui permet notamment de le lier à des centres de soins et à des postes de médecins. On s'intéresse particulièrement à la problématique de décision sur l'état d'un patient afin de détecter d'éventuelles évolutions critiques à long terme. Les données multicapteurs sont exploitées et synthétisées pour représenter la situation médicale d'une personne, modéliser son comportement habituel et la situer à tout instant par rapport à ce modèle en vue d'une décision sur son évolution. Un principe de décision fondé sur les chaînes de Markov est identifié puis expérimenté.

Mots clés : prise de décision, analyse statistique de données, modélisation comportementale, détection d'anomalies, automate fini d'états, chaînes de Markov

Abstract

The development of medical remote care applications is crucial due to the general ageing of the population and the likelihood of a doctors' shortage in the future. This document deals with the smart home for health care concept. This consists in fitting out a flat with multiple sensors and analysing their output signals closely along with other available information such as clinical data. The aim is to detect unusual situations that might give serious cause for concern. The smart home is connected to a large medical network which provides a direct link to medical assistance in case of emergency. We are particularly concerned with making a decision about the patient state so as to detect long term critical evolutions. Data from the multisensors are analysed and synthesized to characterise the patient's state of health, to model his usual behaviour and to decide at any time about the seriousness of the current condition compared with the model. A decision making process based on Markov chains is defined and tested.

Keywords : decision making, statistical data analysis, behaviour modelling, anomaly detection, finite states automata, Markov chains

SOMMAIRE

<u>1</u>	<u>INTRODUCTION.....</u>	<u>6</u>
<u>2</u>	<u>POSITIONNEMENT DU SUJET</u>	<u>7</u>
<u>2.1</u>	<u>TÉLÉMÉDECINE</u>	<u>7</u>
<u>2.2</u>	<u>LA TÉLÉSURVEILLANCE MÉDICALE : L'UNE DES DIMENSIONS DE LA TÉLÉMÉDECINE</u>	<u>7</u>
<u>2.3</u>	<u>DESCRIPTION GÉNÉRALE DU PROJET HIS</u>	<u>7</u>
<u>2.4</u>	<u>IMPLÉMENTATION DU SYSTÈME D'INFORMATION ET DE COMMUNICATION</u>	<u>7</u>
<u>2.5</u>	<u>LES FONCTIONS DU SYSTÈME SIC-HIS</u>	<u>8</u>
<u>2.6</u>	<u>LA DÉTECTION D'ALARMES</u>	<u>9</u>
<u>2.7</u>	<u>SYNTHÈSE ET DÉMARCHE GLOBALE</u>	<u>10</u>
<u>3</u>	<u>DESCRIPTION DU PROCESSUS DE SYNTHÈSE DE DONNÉES EN VUE DE LA DÉCISION</u>	<u>12</u>
<u>3.1</u>	<u>DÉFINITION GLOBALE DU SYSTÈME : LES GRANDES ÉTAPES DU PROCESSUS</u>	<u>12</u>
<u>3.2</u>	<u>DIVERSITÉ DES INFORMATIONS DISPONIBLES</u>	<u>12</u>
<u>3.3</u>	<u>REPRÉSENTATION DE L'ÉTAT DU PATIENT</u>	<u>14</u>
<u>3.4</u>	<u>MODÈLE DE L'ÉVOLUTION HABITUELLE DU PATIENT</u>	<u>17</u>
<u>3.5</u>	<u>IDENTIFICATION DE LA SITUATION DU PATIENT PAR RAPPORT AU MODÈLE D'ÉVOLUTION</u>	<u>24</u>
<u>3.6</u>	<u>DÉCISION SUR L'ÉTAT DU PATIENT</u>	<u>27</u>
<u>3.7</u>	<u>PRISE EN COMPTE DE LA COMPOSANTE TEMPORELLE</u>	<u>29</u>
<u>3.8</u>	<u>ÉVALUATION DES PERFORMANCES DU SYSTÈME</u>	<u>32</u>
<u>3.9</u>	<u>SYNTHÈSE ET SCHÉMA GÉNÉRAL DU PROCESSUS DE DÉCISION</u>	<u>33</u>
<u>4</u>	<u>APPRENTISSAGE DES PARAMÈTRES DU MODÈLE.....</u>	<u>34</u>
<u>4.1</u>	<u>PARAMÈTRES STRATÉGIQUES DU SYSTÈME DE DÉCISION</u>	<u>34</u>
<u>4.2</u>	<u>DONNÉES NÉCESSAIRES À L'APPRENTISSAGE ET À LA VALIDATION DU MODÈLE</u>	<u>34</u>
<u>4.3</u>	<u>DESCRIPTION DU PROCESSUS D'APPRENTISSAGE DU MODÈLE</u>	<u>35</u>
<u>4.4</u>	<u>SYNTHÈSE ET SCHÉMA GÉNÉRAL DU PROCESSUS D'APPRENTISSAGE</u>	<u>38</u>
<u>5</u>	<u>APPLICATION JAVA™ POUR L'IMPLÉMENTATION.....</u>	<u>40</u>
<u>5.1</u>	<u>DÉFINITION DES FONCTIONNALITÉS</u>	<u>40</u>
<u>5.2</u>	<u>PROCESSUS D'EXÉCUTION DE L'APPLICATION</u>	<u>41</u>
<u>5.3</u>	<u>MODÈLE OBJET</u>	<u>42</u>
<u>6</u>	<u>EXPÉRIMENTATION DU MODÈLE</u>	<u>48</u>
<u>6.1</u>	<u>INTRODUCTION</u>	<u>48</u>
<u>6.2</u>	<u>CONDITIONS EXPÉRIMENTALES</u>	<u>48</u>
<u>6.3</u>	<u>PRÉSENTATION DES RÉSULTATS</u>	<u>51</u>
<u>6.4</u>	<u>COMMENTAIRES</u>	<u>57</u>
<u>7</u>	<u>PERSPECTIVES</u>	<u>59</u>
<u>7.1</u>	<u>COLLECTE DE DONNÉES RELATIVES AU PATIENT</u>	<u>59</u>
<u>7.2</u>	<u>GÉNÉRALISATION DU CODAGE DES VARIABLES D'ÉTAT</u>	<u>60</u>
<u>7.3</u>	<u>SIMPLIFICATION DU MODÈLE</u>	<u>60</u>
<u>7.4</u>	<u>APPRENTISSAGE DES PARAMÈTRES DU MODÈLE</u>	<u>62</u>
<u>7.5</u>	<u>INTERVENTION DES PARAMÈTRES ENVIRONNEMENTAUX</u>	<u>62</u>
<u>7.6</u>	<u>D'AUTRES POINTS DE VUE...</u>	<u>62</u>
<u>8</u>	<u>CONCLUSION</u>	<u>63</u>

<u>BIBLIOGRAPHIE</u>	64
<u>ANNEXE A TÉLÉMÉDECINE</u>	67
<u>A.1 DÉFINITION</u>	67
<u>A.2 LES APPORTS DE LA TÉLÉMÉDECINE</u>	67
<u>A.3 LES FREINS</u>	68
<u>A.4 ENJEUX ET NÉCESSITÉ</u>	69
<u>ANNEXE B DESCRIPTION DE L'APPARTEMENT PROTOTYPE</u>	70
<u>B.1 LES CAPTEURS</u>	70
<u>B.2 TRANSMISSION DES INFORMATIONS DES CAPTEURS</u>	71
<u>ANNEXE C LE SYSTÈME CARDIOVASCULAIRE</u>	73
<u>C.1 DESCRIPTION DU SYSTÈME CARDIOVASCULAIRE</u>	73
<u>C.2 LA PRESSION ARTÉRIELLE</u>	76
<u>C.3 LES RISQUES CARDIOVASCULAIRES</u>	82
<u>C.4 CARACTÉRISATION DES VALEURS DES MESURES CARDIAQUES</u>	85
<u>ANNEXE D L'INDICE DE MASSE CORPORELLE (IMC)</u>	87
<u>D.1 DÉFINITION ET CALCUL</u>	87
<u>D.2 ESTIMATION DE LA CORPULENCE D'UN INDIVIDU</u>	87
<u>D.3 CARACTÉRISATION DE LA VALEUR DE L'IMC</u>	88
<u>ANNEXE E TROUBLES DE LA MICTION</u>	89
<u>E.1 DESCRIPTION</u>	89
<u>E.2 CARACTÉRISATION DES TROUBLES DE LA MICTION</u>	89
<u>ANNEXE F GLOSSAIRE MÉDICAL</u>	90
<u>ANNEXE G VALEUR DIAGNOSTIQUE D'UN SIGNE OU D'UN TEST [12]</u>	92
<u>ANNEXE H MODÈLE OBJET DE L'APPLICATION JAVA™ DÉVELOPPÉE</u>	96
<u>H.1 DÉFINITION DES CONSTANTES</u>	96
<u>H.2 DESCRIPTION DES CLASSES</u>	97
<u>H.3 INTERACTIONS ENTRE LES OBJETS</u>	100
<u>ANNEXE I DONNÉES DISPONIBLES POUR L'EXPÉRIMENTATION</u>	105
<u>I.1 DONNÉES INITIALES</u>	105
<u>I.2 DONNÉES DE TEST</u>	106

1 Introduction

Mon stage de DEA a eu lieu dans le laboratoire TIMC (Techniques de l’Imagerie, de la Modélisation et de la Cognition) de l’IMAG (Institut d’Informatique et Mathématiques Appliquées de Grenoble). Il s’est effectué au sein de l’équipe AFIRM (Acquisition, Fusion d’Information et Réseaux pour la Médecine, en cours de création et dont le responsable est N. NOURY) dans le cadre du projet SIC-HIS (Système d’Information et de Communication de l’Habitat Intelligent pour la Santé) pour lequel cette équipe s’investit depuis quelques années. Il s’agit d’un système complet de télésurveillance médicale de personnes à domicile. Il vise à intégrer sur un même réseau médical plusieurs habitats de patients, des centres de télé-vigilance, des postes de médecins et d’autres membres du corps médical. L’objectif est d’assurer un suivi médical à distance de patients, leur permettant ainsi de conserver une large autonomie en toute sécurité. Leur habitat est spécialement équipé de divers capteurs et d’un ordinateur connecté au réseau.

L’objectif de ce stage est l’étude de la problématique de détection de situations critiques d’une personne évoluant dans un habitat spécialement équipé de divers capteurs. Il s’agit plus particulièrement de définir à partir de l’ensemble des données accessibles un système de décision sur l’état du patient à domicile afin de détecter d’éventuelles évolutions critiques.

Mon travail a consisté dans un premier temps en une large étude bibliographique qui a débuté par la recherche d’articles généraux sur la télémédecine, puis plus particulièrement sur la télésurveillance médicale, tout en me documentant sur l’état d’avancement actuel du projet. Le sujet est alors positionné précisément par rapport à cet environnement de travail, ce qui permet de cibler quelques axes théoriques de recherche pertinents pour la suite de la démarche bibliographique. La problématique du sujet s’inscrit dans le domaine général de l’exploration de données (*data mining*). Ce domaine est ensuite progressivement restreint en fonction des objectifs du projet (décision sur la situation du patient) et de ses contraintes, principalement liées au type d’environnement (environnement médical). Il s’agit enfin d’identifier un domaine théorique adapté au contexte de l’habitat intelligent pour la télésurveillance médicale afin d’aboutir à une première expérimentation d’un modèle de décision.

Dans ce document, on commence par situer le sujet dans l’environnement de recherche (cf. §2). Une analyse plus fine des objectifs et des contraintes permet de décrire précisément le processus de synthèse de données en vue d’une décision sur l’état d’un patient à domicile et d’identifier un modèle de décision (cf. §3). On définit ensuite le processus d’apprentissage des paramètres de ce modèle (cf. §4). Une application Java™ est alors développée pour répondre aux besoins d’une première expérimentation du modèle (cf. §5). La démarche expérimentale et la présentation des résultats obtenus (cf. §6) entraînent des réflexions qui permettent de réorienter les axes de recherche pour une étape suivante dans la démarche de définition et de mise en œuvre du système de décision (cf. §7). Le dernier paragraphe présente les conclusions de ce travail de recherche (cf. §8).

2 Positionnement du sujet

2.1 Télémédecine

La télémédecine représente l'utilisation des Nouvelles Technologies de l'Information et de la Communication (NTIC) dans le secteur médical [1]. Elle médiatise l'acte médical en interposant un outil de communication entre les médecins ou entre un médecin et son patient. Ainsi, elle remet en cause une partie de la pratique médicale, mais représente un enjeu considérable. Elle va permettre l'amélioration des conditions de vie et la prise en charge de milliers de personnes (cf. Annexe A).

2.2 La télésurveillance médicale : l'une des dimensions de la télémédecine

La télémédecine a aujourd'hui trouvé de nombreux champs d'application, et se décline en différents termes [2, 3]. La télésurveillance médicale est une de ses dimensions. Elle consiste en la surveillance médicale à distance de patients qui peuvent être à leur domicile.

La télésurveillance médicale au domicile des personnes représente une alternative momentanée ou durable à l'hospitalisation et au recours aux établissements d'hébergement de longue durée. Le patient n'est alors plus contraint de renoncer à son domicile et à la vie en société. Il conserve une large autonomie dans son environnement social et privatif tout en bénéficiant de services préventifs de santé [4].

2.3 Description générale du projet HIS

Le projet HIS (« Habitat Intelligent pour la Santé ») développé au laboratoire TIMC de l'IMAG s'inscrit dans la problématique de surveillance de patients à domicile. Il développe le concept d'un *habitat intelligent*. Il s'agit d'équiper un logement de divers capteurs (capteurs médicaux, capteurs environnementaux, etc.) gérés par un système informatique qui analyse les signaux en temps réel ou différé en s'appuyant également sur d'autres données de type clinique par exemple (symptômes ou pathologie du patient, antécédents médicaux, etc.). L'objectif est de détecter les besoins du patient, les situations inhabituelles et urgentes (chutes, malaises, hypotension, etc.) nécessitant l'intervention d'un centre de soins.

L'habitat intelligent fait partie d'un projet plus vaste définissant le système d'information complet associé au logement télésurveillé dans son environnement médico-social : SIC-HIS (« Système d'Information et de Communication de l'Habitat Intelligent pour la Santé »). Il intègre sur un même réseau médical plusieurs habitats de patients, des centres de télé-vigilance, des postes de médecins et d'autres membres du corps médical et des postes de consultants occasionnels.

A terme, l'idée est bien sûr de pouvoir gérer simultanément les alertes émanant de plusieurs patients à domicile. Cela implique la définition d'ordres de priorité, de déportation d'alarmes vers des postes secondaires en cas de surcharge du poste principal afin d'assurer la sécurité du système. Dans un premier temps, l'étude est limitée à la surveillance d'un seul habitat.

2.4 Implémentation du Système d'Information et de Communication

Le système SIC-HIS est construit selon une architecture d'objets distribués et implémentés en Java. Il est composé de trois grands sous-ensembles répartis sur trois lieux distincts : la station HIS est située dans l'habitat, la station médicale dans le poste de télésurveillance médicale et les postes clients accèdent au système via un serveur internet/intranet. Chaque sous-ensemble est lui-même divisé en plusieurs modules réalisant chacun une fonction du système : la perception et le pré-traitement des signaux est par exemple réalisée au niveau de la station HIS alors que le traitement des alarmes s'effectue à partir de la station médicale (cf. Figure 1).

Le système complet est représenté par le modèle objet global de l'application. Les différents objets sont développés en Java et communiquent par les couches TCP/IP d'Internet mises en œuvre par le protocole RMI (« Remote Method Invocation ») de Java.

Le schéma général du système SIC-HIS est présenté sur la Figure 1.

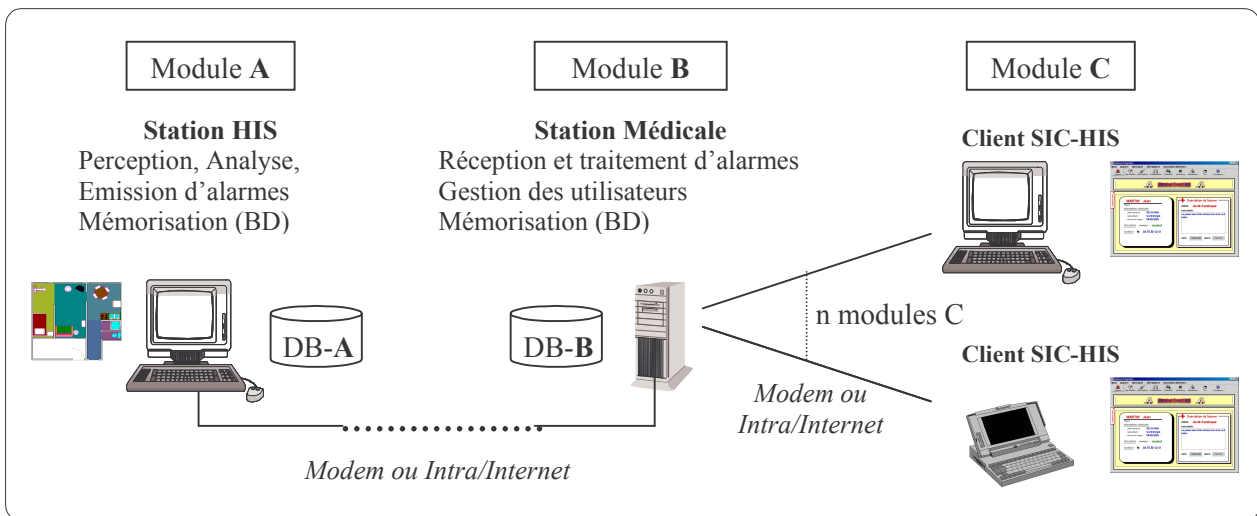


Figure 1. Schéma général du système SIC-HIS

2.5 Les fonctions du système SIC-HIS

Le système SIC-HIS est une chaîne complète de traitement de l'information dont les étapes sont présentées ci-dessous :

- **Perception du patient et de l'environnement** par divers types de capteurs installés dans l'appartement ou bien portés par la personne à domicile.
- **Transmission des informations** par un bus domotique vers un logiciel de traitement des signaux. L'état des capteurs peut être visualisé à tout instant.
- **Gestion du dossier médical** afin d'avoir connaissance de toutes les données cliniques nécessaires au suivi médical du patient. On dispose notamment des fonctionnalités suivantes :
 - Rédaction et gestion d'ordonnances et de comptes-rendus de visite (« Mémos ») ;
 - Droits d'accès et niveau de confidentialité ;
 - Scénarios d'événements communs ou à risque.
- **Analyse et interprétation des données du système** : cet ensemble de données se compose des signaux reçus des capteurs en temps réel et de leur évolution temporelle, ainsi que d'autres données connues du système (dossier médical du patient). Il s'agit de fusionner les données dont on dispose afin d'avoir à tout instant une représentation de l'état global de la personne à domicile en terme médical. L'objectif est la détection en temps réel de tout événement critique pour le patient (chute ou crise cardiaque par exemple). Il s'agit également d'identifier toute détérioration de l'état de la personne sur un plus long terme, pouvant traduire la nécessité de soins ou bien d'une hospitalisation.
- **Présentation de courbes de tendance relatives au patient** : l'évolution des valeurs des données médicales dont on dispose pour représenter l'état de la personne (poids, tension par exemple) est affichée sous la forme conviviale de courbes ou d'histogrammes. Le médecin traitant dispose ainsi à distance de données lui permettant d'émettre un diagnostic sur l'état de la personne à domicile.
- **Déclenchement de différents types d'alarmes suivant les situations détectées** : les événements et les alarmes déclenchées sont enregistrés dans la base de données du système HIS. Les alarmes sont envoyées vers un poste de télésurveillance qui en assure leur traitement.

Dans la suite du document, on s'intéresse à la fonctionnalité d'analyse et d'interprétation des données du système afin de détecter les situations critiques de la personne à domicile

2.6 La détection d'alarmes

La détection d'alarmes est réalisée au niveau de la station HIS (« Habitat Intelligent pour la Santé »), située dans l'habitat de la personne surveillée à distance. L'objectif est la détection automatique des situations à risque pour le patient et l'envoi d'une alarme vers le poste de télésurveillance auquel l'habitat est relié.

Les données reçues des différents capteurs installés dans l'habitat ou embarqués sur la personne sont analysées en temps réel. L'objectif de cette analyse est de prendre une décision sur la nature de l'état global de la personne à domicile afin de détecter toute situation anormale. C'est un problème de synthèse de données tenant compte à la fois des valeurs instantanées délivrées par les capteurs ainsi que de l'évolution temporelle de ces valeurs en vue d'une décision sur la situation du patient. D'autres types d'information pourront également être pris en compte dans le processus de décision :

- données cliniques : elles peuvent par exemple être utiles à la définition de plages de normalité des différents paramètres médicaux pour la personne télésurveillée, ou bien être directement considérées comme des paramètres du processus de détection.
- données obtenues par questionnaire régulier de la personne à domicile : elles peuvent par exemple être le reflet de l'opinion de la personne elle-même sur son état.
- données issues de l'apprentissage : elles caractérisent les habitudes comportementales et cliniques de la personne à domicile.

On distingue deux grands types d'alarmes correspondant à deux processus distincts de détection de situations critiques :

- Détection immédiate d'alarmes urgentes :

Il s'agit d'alarmes déclenchées suite à la réception de données directement interprétables comme étant alarmantes pour le patient (chute grave ou arrêt cardiaque par exemple). Cette détection est la conséquence de la reconnaissance de scénarios ou valeurs critiques d'événements connus.

Un événement est généré par l'analyse combinée des valeurs délivrées par les capteurs et des données du système. Cet événement correspond donc à un nouveau paramètre défini par fusion des paramètres reçus des différents capteurs et/ou connus du système (événement « crise cardiaque » par exemple si on détecte une chute et l'absence de rythme cardiaque). Certains événements peuvent être définis directement par une donnée fournie par un capteur (événement « chute » par exemple).

- Détection à long terme de situations critiques :

Ce type de détection résulte de l'observation de la personne sur une longue période initialisée au moment de l'installation du système dans l'appartement. Il s'agit de détecter toute dégradation lente de l'état du patient à partir de l'observation de l'évolution des données environnementales et médicales reçues en temps réel par les capteurs (poids, tension, etc.) et des informations cliniques dont on dispose.

Une base de données d'événements liée au système HIS permet de garder une trace des événements détectés et des alarmes envoyées au centre de télésurveillance.

Ce sont les détections d'alarmes sur le long terme qui nous préoccupent dans ce sujet. Les paragraphes suivants traitent donc de cet aspect de la problématique de synthèse de données.

2.7 Synthèse et démarche globale

Les paragraphes précédents permettent de situer le sujet de recherche par rapport au contexte de travail allant du cadre général de la télémédecine jusqu'au projet développé au laboratoire TIMC/IMAG. Ce positionnement du sujet est synthétisé sur la Figure 2.

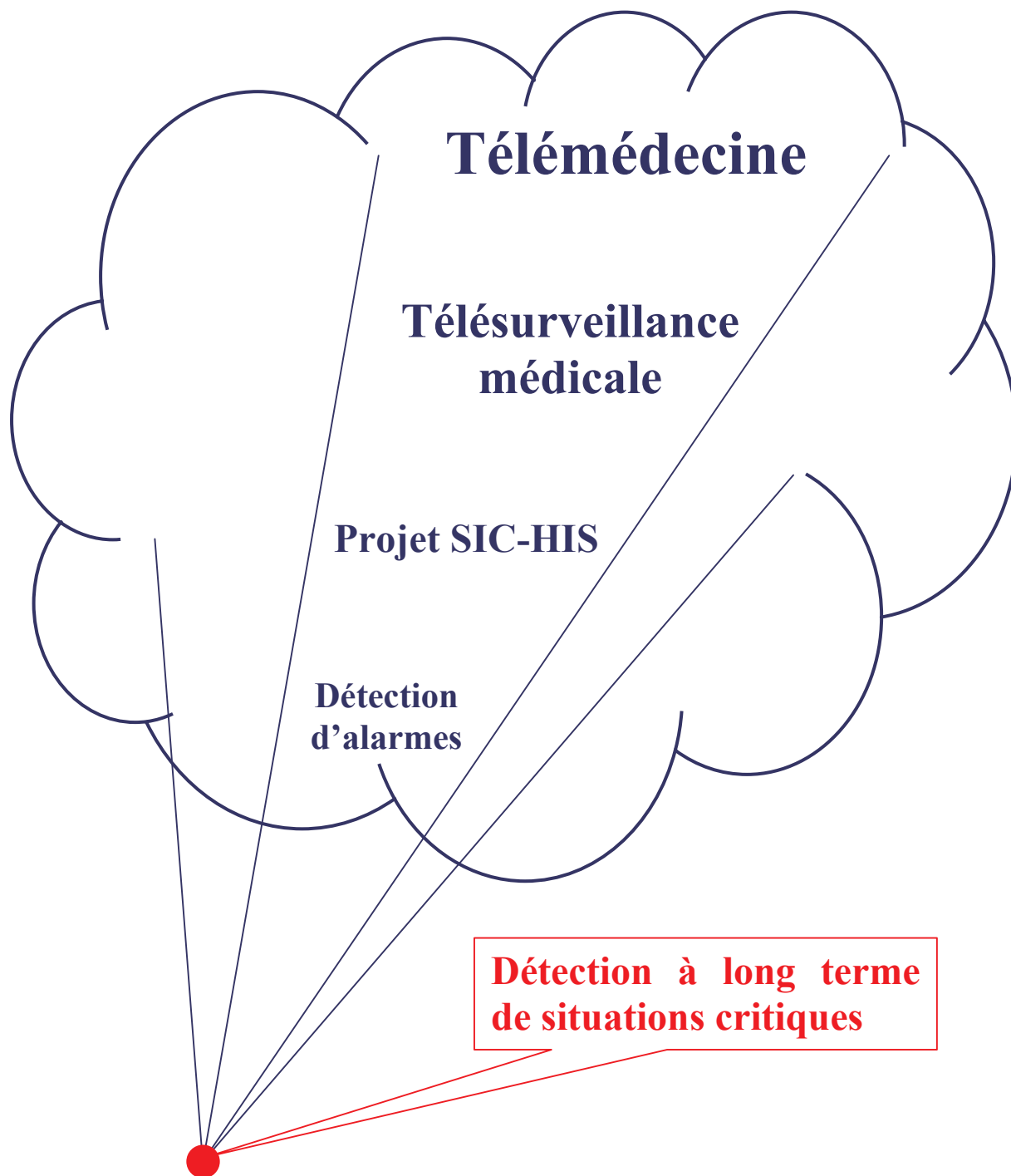


Figure 2. Synthèse du contexte de travail

Afin de bien cadrer la problématique de synthèse de données au sein de l'application SIC-HIS, il convient d'approfondir la définition des objectifs et des contraintes du projet donnée dans le paragraphe précédent de présentation de la problématique de détection d'alarmes (cf. §2.6).

Les caractéristiques du projet et les problèmes soulevés induisent les grandes étapes du fonctionnement du système. Chaque étape met en évidence différentes problématiques dont la résolution passe par une ou plusieurs approches théoriques. Les grands axes de recherche qui se rapprochent de la problématique du sujet doivent être identifiés. Une analyse précise de l'état de l'art dans les différents domaines théoriques concernés est nécessaire pour déterminer les techniques *a priori* les plus adaptées aux objectifs de la recherche. Des solutions théoriques sont alors envisagées puis appliquées au projet SIC-HIS. Il faut enfin inscrire le développement du système dans un cycle théorique d'évaluation/modification assurant sa convergence vers un état de plus en plus performant.

Le schéma de la Figure 3 présente la démarche globale de recherche autour du problème de synthèse de données en vue d'une décision sur l'état de la personne à domicile.

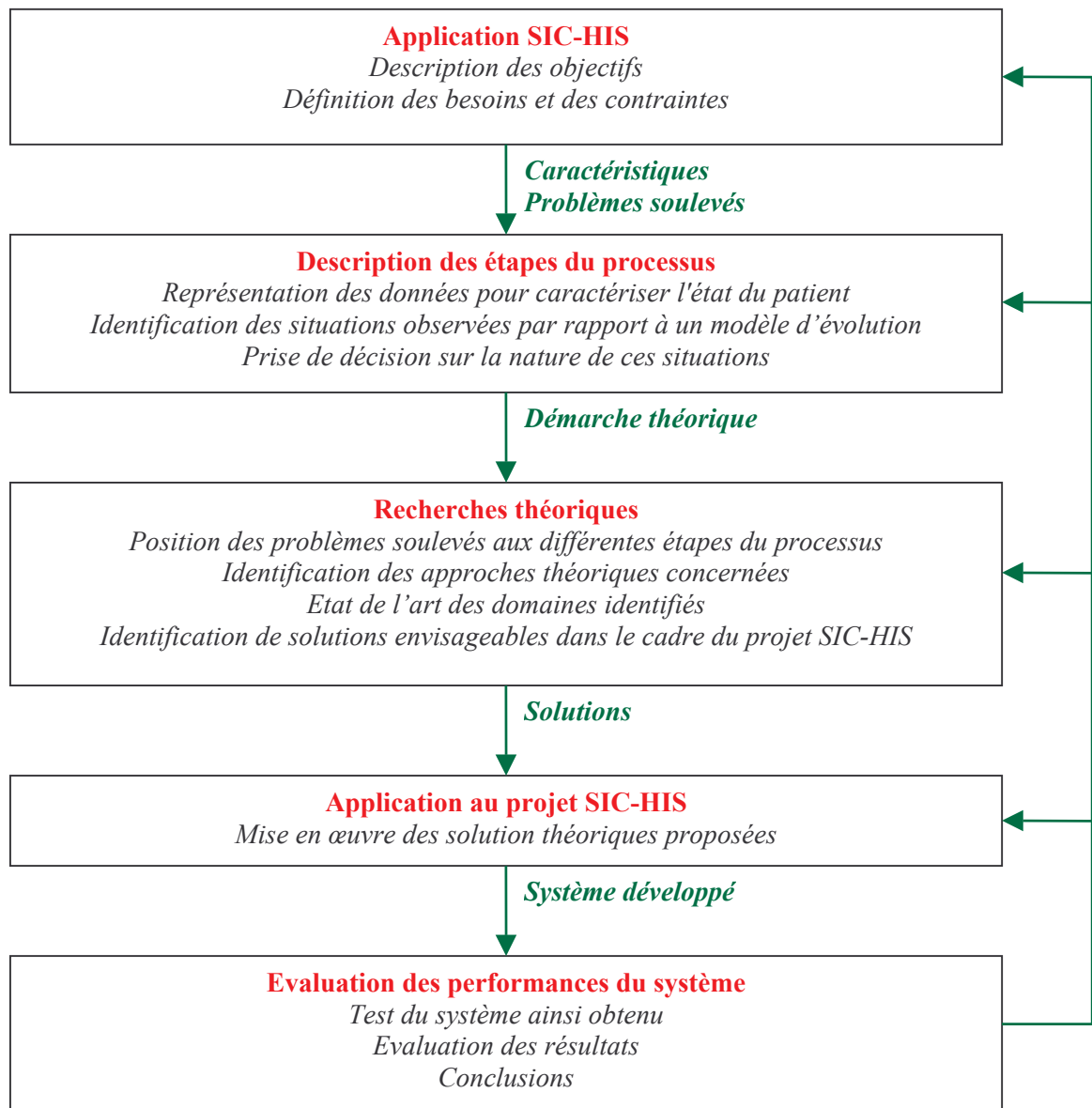


Figure 3. Démarche globale de recherche

3 Description du processus de synthèse de données en vue de la décision

3.1 Définition globale du système : les grandes étapes du processus

Le système de synthèse de données au sein du projet SIC-HIS est un processus qui doit permettre à partir de l'exploration de différents types de données, appelées *observables*, de prendre une décision en temps réel sur l'état de la personne à domicile.

On identifie trois grandes étapes lors du déroulement du processus de détection de situations critiques :

- **La représentation de l'état du patient :**

Les informations disponibles en entrée du système (poids, tension, indicateurs de présence, etc.) sont combinées ou utilisées directement pour définir les valeurs d'un ensemble de variables médicalement significatives reflétant l'état du patient à un instant donné, appelé état courant. Ces variables forment le vecteur d'état du patient.

- **L'identification de cet état par rapport au modèle d'évolution normale du patient :**

On cherche à situer l'état courant par rapport à un modèle du comportement habituel du patient. Il s'agit donc de définir l'écart entre une situation observée et un modèle d'états représentatifs de l'évolution du patient.

- **La décision sur la nature de la situation observée :**

L'état courant est classé parmi un ensemble de situations possibles définies au préalable : ce sont les situations de référence. On peut par exemple décider de diviser les états observés en trois classes : « situation habituelle », « situation préoccupante » et « situation critique ». Cette classification est réalisée à partir de la distance calculée d'un état courant par rapport au modèle d'évolution normale de la personne à domicile. Plus la situation est estimée éloignée du modèle, plus elle est critique.

Le déroulement de ce processus exige donc une **phase d'apprentissage** pendant laquelle on construit un modèle d'évolution de la personne à domicile à partir d'information *a priori* (données de type clinique par exemple) et de données obtenues suite à l'observation de son évolution habituelle dans l'habitat.

Dans les paragraphes suivants sont décrits successivement les types d'informations disponibles, les principes de représentation de l'état du patient, de modélisation des évolutions possibles de ces états, d'identification d'une situation à un instant donné par rapport à ce modèle et de décision sur le type de situation observée. La prise en compte de la composante temporelle ainsi que l'évaluation des performances du système sont également discutés. On présente finalement un schéma général du processus de synthèse de données pour la détection de situations critiques.

3.2 Diversité des informations disponibles

La détection des situations critiques d'une personne à domicile est réalisée à partir d'un large ensemble de données de différente nature et stockées dans la base de données du système SIC-HIS. Il s'agit d'en extraire des informations pertinentes vis à vis de la problématique posée.

Grâce au développement de l'informatique, les systèmes d'extraction de connaissances sont réalisés à partir d'ensembles de données de plus en plus abondants et diversifiés. Pour faire face à ces nouvelles contraintes, l'informatique médicale utilise les technologies développées dans le nouveau domaine de recherche interdisciplinaire d'extraction de connaissances dans les bases de données : *Knowledge Discovery in Database (KDD)* [5 , 6]. Il s'agit tout d'abord de comprendre le contexte de recherche, d'identifier l'ensemble des données à prendre en compte selon la problématique considérée, d'en extraire des régularités non évidentes *a priori* et de les formuler ensuite sous la forme de modèles ou de règles [7, 8].

Dans le contexte médical, l'extraction de connaissances est destinée au diagnostic, à la recherche de pathologies, à la prédiction, à la surveillance, au soutien thérapeutique et au suivi du patient en général

[6]. Les techniques formelles utilisées sont issues des théories de la décision [9] et du raisonnement symbolique [10]. Elles comprennent des méthodes situées à l'intersection de ces deux domaines de recherche comme les réseaux de croyance probabilistes [11]. L'analyse intelligente de données inclut les techniques d'apprentissage, de classification, de représentation de données et d'interprétation de séquences de données [6].

Dans le cadre du projet SIC-HIS, la complexité du problème de synthèse de données réside en premier lieu dans l'hétérogénéité des informations disponibles pour évaluer l'état de la personne à domicile [6]. Ces informations correspondent dans certains cas à des données issues de mesures précises, par exemple pour ce qui concerne les données physiologiques obtenues par des capteurs comme le poids ou la température et dont la valeur est un nombre réel. Mais il peut également s'agir de données qualitatives liées aux particularités du patient : données cliniques rapportées par un médecin et relatives aux symptômes de la personne par exemple.

Il ne s'agit pas de rassembler le plus d'informations possibles de toutes sortes. Il faut également veiller à leur pertinence, c'est-à-dire à leur adéquation à la finalité du système. La prise en compte d'une information non pertinente n'apporte aucune information utile par rapport à la problématique de détection des situations critiques de la personne à domicile. Cette information complètement indépendante du problème posé ne fait qu'ajouter de la complexité et du bruit dans le système, diminuant ainsi sa fiabilité.

On désigne par le terme « observables » l'ensemble des informations disponibles pour la télésurveillance du patient à domicile. Ces informations sont symbolisées par un ensemble de m données accessibles à un tout instant t : $\{Y_{t,i}\}_{0 < i \leq m}$. A un instant t donné, l'ensemble des observables est donc représenté par le vecteur $\mathbf{Y}_t = [Y_{t,1}, Y_{t,2}, \dots, Y_{t,m}]$. Chaque composante $Y_{t,i}$ correspond à une donnée brute délivrée par un capteur (poids ou indicateur de présence dans la salle de bain par exemple) ou bien à une donnée clinique du patient (âge ou sexe par exemple).

Ces informations disponibles sont divisées en trois catégories définissant trois types de provenance possibles :

- [les informations perceptives](#) : en évolution permanente, elles correspondent aux données délivrées en temps réel par les capteurs ;
- [les informations méta-perceptives](#) : elles représentent un historique de données relatives au patient qui peuvent prendre la forme d'un modèle d'évolution ou d'un ensemble de situations de référence par exemple ;
- [les informations cliniques](#) : elles correspondent aux données du dossier médical du patient. Elles sont stockées dans la base de données liée à l'habitat télésurveillé et sont peu modifiées.

Ces trois types d'information sont décrits dans les paragraphes ci-après.

3.2.1 Les informations perceptives

Les capteurs disposés dans l'appartement délivrent en temps réel des données relatives au patient comme la tension, le poids, ou encore des indicateurs de présence dans chacune des pièces. Ces données sont envoyées au système dès qu'une nouvelle valeur est disponible : lorsqu'une prise de tension est effectuée par exemple, ou bien lorsque le patient change de pièce dans l'appartement, etc. Chaque nouvelle valeur reçue permet de réévaluer l'état de la personne à domicile et de poser de nouveau la question concernant sa nature : la nouvelle situation ainsi définie est-elle critique ou non pour le patient ?

La prise en compte de ces informations pose le problème de la gestion de l'incertain. La présence d'imprécisions et d'incertitudes au niveau des *observables* (c'est-à-dire des données disponibles en entrée du système) peut avoir plusieurs origines liées à la nature et à la répartition des sources d'information [6]:

- [Incomplétude de l'information](#) : dans le cas par exemple où le patient manque de rigueur dans les prises de mesures qu'il doit effectuer régulièrement (poids, température, tension par exemple). Il manque alors en entrée du système certaines informations pourtant nécessaires à l'évaluation de l'état de la personne à domicile.

- **Imprécision de l'information** : elle concerne également les données issues de mesures effectuées par la personne à domicile elle-même. Une tension prise dans de mauvaises conditions, juste après une grande activité par exemple, n'est alors plus le reflet fiable de l'état cardiaque du patient.
- **Incohérences entre des sources d'information** : certaines données comme la fréquence cardiaque sont accessibles à partir de différents capteurs : le tensiomètre et l'actimètre (cf. Annexe B). Les informations reçues de ces deux sources doivent alors être cohérentes, c'est-à-dire qu'on s'attend à ce que les mesures issues de ces capteurs à des instants voisins soient identiques à une faible marge d'erreur près. C'est la redondance des informations qui permet de détecter ces incohérences.
- **Manque de fiabilité des sources** : la panne d'un capteur par exemple peut entraîner la prise en compte de données aberrantes et non significatives pour le patient. Ces données non valides doivent être écartées du processus de décision. La validité d'une information peut être estimée par un degré de confiance en la source, une confiance nulle représentant alors l'état d'ignorance.

Les données obtenues via les capteurs définissent un ensemble de variables à valeur soit dans un ensemble discret (l'ensemble {Oui, Non} par exemple pour l'indicateur de présence dans une des pièces de l'appartement), soit dans un ensemble continu (l'ensemble des réels positifs par exemple pour le poids). La définition des variables d'état fait apparaître la notion de flou, liée à l'ambiguïté au niveau de l'interprétation des valeurs qu'elles peuvent prendre. Une valeur de poids par exemple n'a aucune signification dans l'absolu mais uniquement relativement à un ensemble de valeurs habituelles ou jugées normales pour un patient donné. De plus, il n'y a pas de vocabulaire standardisé communément utilisé par la communauté médicale et répondant à des définitions clairement exprimées [12].

3.2.2 Les informations méta-perceptives

Ces informations correspondent à un ensemble de valeurs habituelles prises par les différentes variables physiologiques dont la valeur est obtenue à un instant donné par les capteurs (évolutions normales du poids, de la tension, etc.).

La connaissance pour un patient de l'évolution habituelle des variables physiologiques délivrées par les capteurs (poids, tension, etc.) est notamment importante au niveau de l'interprétation des valeurs prises par ces variables. Chacune d'entre elles est donc caractérisée par un ensemble de séquences de valeurs représentatives d'un comportement considéré comme normal pour le patient. Ces ensembles de données sont utiles à la représentation de l'état de la personne à domicile ainsi qu'à la construction d'un modèle d'évolution.

3.2.3 Les informations cliniques

Ces informations correspondent aux données du dossier médical du patient et représentent donc son état clinique.

Le dossier médical du patient donne accès à un ensemble de données cliniques : âge, sexe, antécédents médicaux par exemple. Ces données sont stockées dans la base de données du système HIS de l'habitat télésurveillé. Elles peuvent être consultées et éventuellement modifiées à partir d'un poste relié au réseau médical du SIC-HIS par les utilisateurs autorisés (médecins, personnel médical, parents).

3.3 Représentation de l'état du patient

3.3.1 Variables d'état du patient

On cherche à définir une représentation mathématique médicalement significative de l'état du patient à partir des différentes informations dont on dispose (données issues des capteurs, données cliniques) et de leurs caractéristiques (problèmes d'hétérogénéité des données et de gestion de l'incertain). Il s'agit de combiner ces données de nature différente pour représenter la situation de la personne à domicile. Dans le cadre de la présente étude, cette étape de fusion de données n'est pas particulièrement abordée. Elle est

cependant nécessaire à la définition de variables synthétiques de qualité (variables d'état) représentant l'ensemble des observables.

L'objectif de cette étape est ainsi de définir un ensemble de n variables d'état, $\{X_{t,i}\}_{0 < i \leq n}$, ayant une signification précise en terme médical comme le poids, la tension, la distance moyenne de déplacement au cours d'une journée ou encore la pollakiurie (cf. Annexe E). Ce sont ces variables qui, à un instant t donné, forment le vecteur d'état du patient : $\mathbf{X}_t = [X_{t,1}, X_{t,2}, \dots, X_{t,n}]$. Chaque variable $X_{t,i}$ a une signification précise en terme médical et doit être pertinente pour la problématique de détection des situations à risque de la personne à domicile.

Les variables d'état du patient sont obtenues à partir des observables. On est donc amené à définir une fonction de représentation permettant de passer du niveau des données brutes, $\{Y_{t,i}\}_{0 < i \leq m}$, au premier niveau interprétatif : $\{X_{t,i}\}_{0 < i \leq n}$.

3.3.2 Fonction de représentation

Le principe de représentation de l'état du patient consiste à définir une fonction de l'espace des observables, $\{Y_{t,i}\}_{0 < i \leq m}$, constitué des informations brutes dont on dispose (poids, tension, indicateurs de présence, etc.), dans l'espace des variables d'état, $\{X_{t,i}\}_{0 < i \leq n}$, formé d'un ensemble de variables médicalement significatives (poids, tension, pollakiurie, distance journalière parcourue, etc.). Parmi les exemples cités précédemment, le poids et la tension sont des variables d'état ($X_{t,1}$ et $X_{t,2}$ par exemple) dont les valeurs sont obtenues directement à partir des capteurs, c'est-à-dire à partir des observables ($Y_{t,1}$ et $Y_{t,2}$ par exemple). Dans ce cas particuliers, les observables $Y_{t,1}$ et $Y_{t,2}$ sont directement interprétables comme des variables d'état du patient : $X_{t,1}$ et $X_{t,2}$. La fonction de représentation est alors la fonction identité :

$$\text{Pour } i \in \{1, 2\}, X_{t,i} = Y_{t,i}$$

Par contre, la distance moyenne de déplacement de la personne est le résultat de la combinaison de plusieurs informations : on peut par exemple la calculer à partir des successions de notifications de présence dans chacune des pièces qui donnent une idée des déplacements de la personne dans l'appartement et en tenant compte de l'âge du patient. Dans cet autre cas il faut donc combiner plusieurs données de natures différentes et la fonction de représentation n'est plus si simple :

$$\exists i, 0 < i \leq n, \text{ tel que } X_{t,i} = f(Y_{t,1}, \dots, Y_{t,j}, \dots, Y_{t,m})$$

3.3.3 Codage des variables d'état : une représentation du patient orientée vers la modélisation

Ce paragraphe met en évidence la nécessité de qualifier parmi un ensemble fini de possibilités les valeurs prises par les variables d'état pour la modélisation de l'évolution de la personne à domicile.

Certains acteurs du système SIC-HIS ont accès en permanence à des courbes de tendances représentant l'évolution des variables d'état. Il est alors intéressant de leur présenter les valeurs les plus précises dont on dispose pour ces variables. Ils savent alors les interpréter en fonction de leurs connaissances sur le patient. On trace par exemple la courbe d'évolution du poids à partir des mesures qui ont été prises à différents instants par le patient.

Cependant, il ne faut pas perdre de vue que la représentation des données est également destinée à l'identification des situations observées afin de détecter celles qui sont inhabituelles pour la personne à domicile. Cette reconnaissance de situation est effectuée par rapport à un ensemble fini d'états de référence. Pour obtenir un nombre fini d'états possibles pour le patient et en limiter le nombre, il est primordial de qualifier de manière assez globale la valeur prise par chaque variable d'état du patient. En reprenant l'exemple de la variable d'état « poids », il est surtout important au niveau du système de détection des situations critiques de savoir si ce poids est habituel, ou bien plutôt diminué ou augmenté.

Afin de disposer d'une représentation de l'état du patient homogène, significative en terme médical et plus facilement exploitable au niveau du système de décision, une approche très intéressante est de coder l'ensemble des valeurs possibles des variables d'état à valeur dans un ensemble continu afin de les qualifier en terme médical. L'intervalle des valeurs prises par les variables continues médicalement significatives comme le poids peut par exemple être divisé en trois intervalles correspondant aux valeurs

« habituelles », « diminuées » et « augmentées ». Il convient alors de déterminer pour chaque variable d'état continue le nombre d'intervalles à considérer et leurs délimitations.

L'étape de codage des variables d'état est fortement liée à la problématique suivante qui consiste à modéliser l'évolution habituelle des états possibles du patient. En effet, dans le cas où on considère un graphe fini d'états composé de l'ensemble des vecteurs d'état possibles du patient, le nombre d'états considérés est directement dépendant du nombre d'intervalles définis pour chaque variable d'état. On considère par exemple disposer uniquement de trois variables d'état continues : le poids, la fréquence cardiaque et la température. On segmente l'ensemble des valeurs possibles de ces variables d'état en trois intervalles : « diminué », « habituel » et « augmenté ». On obtient ainsi $N = 3^3 = 27$ états possibles pour le patient. Si on ajoute maintenant un intervalle correspondant aux valeurs qualifiées de « très augmentées » pour une des variables (le poids par exemple), on augmente le nombre d'états à considérer qui s'élève alors à $N' = 3^2 \times 4 = 36$ états.

3.3.4 Intervalles de variation des variables d'état

Pour une variable d'état du patient donnée, on cherche à définir un intervalle de variation correspondant aux fluctuations habituelles des valeurs de cette variable. Cet intervalle correspond à l'ensemble des valeurs considérées comme « normales » pour le patient. En effet, selon les caractéristiques de la personne télésurveillée, l'intervalle de variation habituel est inclus ou non dans l'intervalle des variations considérées comme normales par la profession médicale pour une variable d'état donnée. Un patient légèrement hypertendu a par exemple un intervalle de variation de sa pression artérielle diastolique légèrement au-dessus de l'intervalle de variation considéré comme normal par la profession médicale (cf. Annexe C). Cet exemple est illustré par le schéma de la Figure 4.

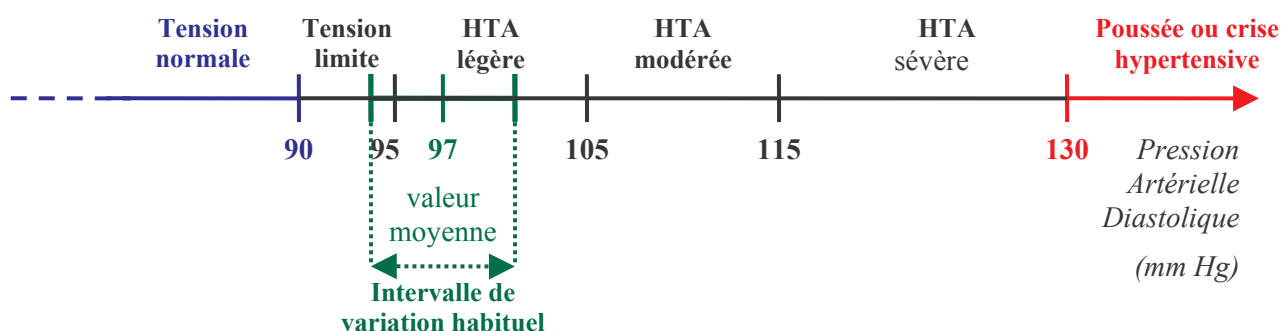


Figure 4. Répartition des valeurs de la pression artérielle diastolique d'un patient hypertendu par rapport aux valeurs limites définies par la profession médicale

Il faut donc bien distinguer pour chaque variable d'état deux types d'intervalles de variation :

- l'intervalle « normal » de variation, correspondant aux limites de variation définies usuellement en médecine : [18.5, 25] pour l'Indice de Masse Corporelle (IMC) par exemple (cf. Annexe D) ;
- l'intervalle « habituel » de variation qui est l'intervalle de variation de chaque variable d'état pour un patient donné. Cet intervalle est défini suite à l'observation de l'évolution du patient, c'est-à-dire à la mesure de l'ensemble des valeurs des variables d'état à différents instants. Cet intervalle est considéré comme « normal » pour un patient donné.

En dehors de l'intervalle habituel de variations, les valeurs supérieures seront considérées comme « augmentées » et les valeurs inférieures comme « diminuées ». On segmente ainsi l'ensemble des valeurs possibles de chaque variable d'état continue en trois intervalles disjoints correspondant aux valeurs considérées comme « diminuées », « habituelles », puis « augmentées ». La délimitation de ces intervalles de variation peut être réalisée soit manuellement à partir des connaissances des experts en médecine, soit en appliquant des algorithmes d'apprentissage.

Les connaissances des experts en médecine permettent la définition d'intervalles de confiance autour d'une valeur moyenne habituelle de chaque variable d'état continue pour le patient [13, 14].

Pour la détermination automatique des intervalles de segmentation des valeurs des variables d'état, il est nécessaire de disposer d'un ensemble de valeurs représentatives de l'évolution habituelle du patient. On considère par exemple un ensemble de m valeurs $\{x_{0,i}, x_{1,i}, x_{2,i}, \dots, x_{m-1,i}\}$ pour la variable d'état X_i . Deux mesures sont utiles à la délimitation de l'intervalle des valeurs habituelles de cette variable : la valeur moyenne m_i et l'écart-type σ_i .

Si on considère l'ensemble de n valeurs $\{x_0, x_1, x_2, \dots, x_{n-1}\}$ pour la variable aléatoire X , on calcule à partir de cet échantillon la moyenne m et la variance σ^2 par les formules suivantes [15]:

$$m = \frac{1}{n} \sum_{i=0}^{n-1} x_i \quad \text{et} \quad \sigma^2 = \frac{1}{n} \sum_{i=0}^{n-1} (x_i - m)^2$$

L'écart-type σ est la racine carrée de la variance. Il s'exprime dans la même unité que la variable étudiée X .

L'intervalle « habituel » de variations peut alors être défini simplement comme l'intervalle centré sur la valeur moyenne m_i et ayant une largeur de deux écart-types par exemple : $[m_i - \sigma_i, m_i + \sigma_i]$.

Le problème de la délimitation stricte d'un intervalle considéré comme « normal » pour le patient est qu'un changement mineur de la valeur d'une des variables d'état peut avoir un effet drastique sur la représentation de la situation si la nouvelle valeur sort de l'intervalle habituel de variation défini pour cette variable. Une solution serait de considérer plutôt des intervalles de segmentation flous [6, 13, 16]. L'introduction de la notion d'ensembles flous semble approprié aux problèmes de décision dans le domaine de la médecine puisque les connaissances médicales et les décisions qui en résultent sont plus ou moins incertaines [17, 18]. Les connaissances cliniques sont en effet l'expression d'observations statistiques sur le patient [12].

Les intervalles de segmentation des variables d'état doivent être révisés périodiquement pour prendre en compte les évolutions naturelles des valeurs habituelles prises par les différentes variables. En vieillissant, un patient a tendance à avoir un poids plus important ou bien une fréquence cardiaque plus élevée par exemple. Ces constatations induisent un décalage dans le temps des intervalles qui avaient été définis, voire une modification de la largeur de chacun d'entre eux. Le processus de décision sur la situation d'une personne à domicile est bien dynamique et doit donc prendre explicitement en compte la notion du temps. Cette remarque met également en évidence le fait qu'un écart par rapport au modèle d'évolution de la personne à domicile n'est pas forcément inquiétant puisqu'il peut s'agir d'une évolution naturelle avec l'âge et le contexte de vie. Il faut donc savoir interpréter ces écarts pour en déterminer la gravité ou non.

3.4 Modèle de l'évolution habituelle du patient

La connaissance de données représentatives de l'évolution habituelle du patient permet de construire un modèle d'évolution de la personne à domicile. Ce modèle est la base de la décision sur la situation du patient à tout instant.

3.4.1 Les approches bibliographiques de la décision médicale

La prise de décision dans un contexte médical est effectuée à partir d'un ensemble de données hétérogènes et incertaines. Il s'agit d'analyser ces données afin d'en extraire les informations utiles par rapport au problème de décision considéré. Les techniques d'apprentissage (*machine learning*) [10] sont particulièrement utilisées pour l'analyse intelligente de données en médecine [19]. Leur objectif est d'améliorer la prise de décision médicale par l'automatisation de l'analyse des bases de données relatives aux patients en tenant compte des expériences passées [6].

Les méthodes liées à l'apprentissage sont classées en trois principaux groupes [6, 13] :

- l'apprentissage de règles symboliques par des méthodes d'induction : l'induction de règles, (*induction of rules*) [20], les arbres de décision (*decision trees*) [21] et la programmation logique (*logic programs*) [22] ;
- les méthodes statistiques ou de reconnaissance des formes : la méthode des K-plus proches voisins [23], l'apprentissage basé sur l'exemple (*instance-based learning*) [24], l'analyse discriminante, les classifieurs Bayésiens ;
- les réseaux de neurones artificiels (*ANN : Artificial Neural Networks*) : *neural networks with back-propagation learning*, *Kohonen's self-organizing network*, *Hopfield's associative memory*.

Le choix d'une méthode d'analyse de données dépend des particularités de l'environnement de décision. Dans le contexte médical, il est important de pouvoir expliquer et argumenter les choix d'un système de décision. Les méthodes de type « boîte noire » (ANN) sont ainsi écartées de cet environnement de recherche. L'analyse descendante (*top-down decision tree learner*) et les classifieurs Bayésiens sont par contre des techniques qui ont fait leurs preuves dans le domaine du diagnostic médical [13].

La résolution d'un problème de décision passe par la traduction de facteurs de décision complexes en un modèle paramétré simple dépendant d'un ensemble d'hypothèses et de contraintes relatives au contexte [25]. Il convient de choisir le type de modèle utilisé, de paramétrer ce modèle en fonction du contexte (décisions possibles, variables d'état, probabilités de transitions entre états), puis de définir une politique de résolution du problème de décision par rapport à ce modèle [26].

Les caractéristiques de l'environnement médical induisent un ensemble de contraintes à prendre en compte impérativement dans la spécification d'un système de décision [6]. On a déjà noté l'importance de l'interprétabilité d'une analyse de données dans ce contexte. Il faut également tenir compte des caractéristiques intrinsèques des données médicales (*incompleteness, incorrectness, sparseness, exactness*) [27, 28], de l'hétérogénéité des données, des contraintes liées au temps, des objectifs de spécificité (réduction du taux de faux-positifs, c'est-à-dire de fausses détections) et de sensibilité (réduction du taux de faux-négatifs, c'est-à-dire de non-détections) d'une décision (cf. Annexe G).

La problématique de décision dans le domaine médical est ainsi caractérisée par la prise en compte d'informations dynamiques et incertaines [25] : il s'agit d'une analyse décisionnelle dynamique. L'analyse décisionnelle correspond à la problématique de prise de décision dans un contexte incertain [26, 29]. Lorsque la composante temporelle doit également être prise explicitement en compte on parle d'analyse décisionnelle dynamique. Des formalismes spécifiques ont été développés pour répondre à cette contrainte temporelle : *Dynamic influence diagrams, Markov cycle trees, stochastic trees* [26]. Ils sont basés sur les modèles conventionnels de décision que sont les arbres de décision et les diagrammes d'influence, et reposent sur la théorie mathématique des processus de décision semi-markoviens (*SMDP : Semi-Markov Decision Process*). Cette théorie est plus générale que celle des processus de décision markoviens (*MDP : Markov Decision Process*) puisqu'elle prend en plus explicitement en compte le facteur temps [25].

Un processus de décision markovien (MDP) [30, 31] est un processus stochastique qui satisfait à la propriété de Markov (cf. p18) [32]. Il devient un problème de décision markovien s'il est associé à un ensemble de critères de performance [33]. Les MDPs sont exploités dans un large domaine d'applications. Ils sont à la base d'un grand nombre de problèmes étudiés dans le domaine de l'Intelligence Artificielle (IA) : planification automatique de différentes tâches, apprentissage, prises de décision [33]. Un processus de décision markovien se compose de cinq éléments [14, 34] : un ensemble d'états, des probabilités de transition entre les états, un ensemble d'actions, des instants de décision, une fonction d'estimation du coût d'une décision. Dans le cas où le processus est discret, une chaîne de Markov (ensemble fini d'états et de transitions entre états) est sous-jacente à chaque problème de décision markovien.

Les processus de décision semi-markovien (SMDP), plus généraux que les MDPs, sont à la base de la résolution des processus de décision dynamiques, c'est-à-dire prenant explicitement en compte la composante temporelle. La matrice de transition entre états devient ainsi une fonction du temps, la distribution des instants de décision ainsi que les coûts des décisions le deviennent également [26]. Le modèle théorique fondamental des SMDPs est ensuite simplifié en fonction des caractéristiques du

problème considéré, donnant ainsi naissance à différents formalismes tels que *Dynamic influence diagrams*, *Markov cycle trees*, *stochastic trees* [25]. Selon le type de problème posé, il est possible de résoudre un problème de décision semi-markovien par l'un des algorithmes suivant [26, 33] : *value iteration*, *policy iteration*, *adaptive aggregation*, *linear programming*, ou bien considérer directement le cadre théorique à l'origine du modèle pour prendre une décision.

De nombreuses extensions des processus de décision semi-markoviens ont été développées pour répondre aux contraintes particulières de certains projets. Il s'agit par exemple de limiter le nombre de décisions accessibles à chaque étape parmi l'ensemble des décisions possibles, ou encore de procéder à un instant donné à une évaluation partielle du modèle [26]. Les mécanismes d'abstraction permettent de prendre une décision en prenant un compte différents niveaux de détails dans le modèle. Ils sont par exemple particulièrement utiles lorsqu'il existe de fortes contraintes temporelles ou bien lorsque des données manquantes ne permettent plus de définir suffisamment précisément un modèle plus détaillé [35]. Certains formalismes permettent aussi de modéliser un problème de décision de plusieurs points de vue : *transition view*, *influence view*, *tree view* [25, 35]. L'analyse conjointe des différents points de vue permet d'améliorer le modèle et la prise de décision.

L'analyse précédente met en évidence l'efficacité souvent constatée d'une modélisation basée sur les chaînes de Markov pour traiter un problème de décision dans le domaine médical. La théorie mathématique des chaînes de Markov est le fondement de nombreux principes de décision comme les processus de décisions markoviens et semi-markoviens, les chaînes de Markov cachées (*HMM*, *Hidden Markov Decision Process*), les *POMDPs* (*Partially Observable Markov Decision Process*), etc. La chaîne de Markov est alors sous-jacente au problème de décision.

Dans le cadre de la télésurveillance médicale, il s'agit de diagnostiquer l'état d'une personne à domicile à partir d'un ensemble de données hétérogènes issues des capteurs situés dans l'habitat et des données cliniques du patient (dossier médical). S'il est possible de caractériser le comportement d'un patient par un ensemble fini d'états parmi lesquels la personne évolue, on peut alors probablement modéliser son comportement par une chaîne de Markov : un ensemble fini d'états, les probabilités initiales de chacun des états et les probabilités de transition entre états. L'état d'une personne à domicile est défini à tout instant par la donnée d'un ensemble de variables physiologiques et environnementales. La qualification globale des valeurs de ces variables définit un ensemble d'états accessibles par le patient qui peuvent même être qualifiés de façon plus générale comme étant habituel ou bien préoccupant, voire critique.

Le principe de base de définition d'une chaîne de Markov semble ainsi intuitivement adapté à la problématique de détection d'alarmes dans le cadre de la télésurveillance médicale à domicile. Les contraintes et les objectifs du projet permettent d'identifier les types de modèles envisageables, c'est-à-dire l'ensemble des états considérés dans la chaîne de Markov sous-jacente au problème de décision. Selon le type de modèle, il faut ensuite définir les méthodes d'apprentissage des paramètres en fonction des données dont on dispose ainsi qu'un principe de résolution permettant d'aboutir à une prise de décision quant à la nature de la situation du patient.

3.4.2 Détecter une situation critique connue ou bien une situation inhabituelle ?

Les techniques de modélisation puis de détection des situations critiques du patient à domicile peuvent être classées en deux catégories selon le type de modèle utilisé pour identifier l'état courant du patient. Soit on modélise les scénarios d'événements critiques susceptibles de se produire et on examine à partir de la situation courante observée leur possibilité d'occurrence. La question posée est alors : la situation observée s'apparente-t-elle à une des situations critiques modélisées ? Ou bien on modélise l'évolution habituelle de l'état du patient et on essaie de déterminer si l'écart entre une situation observée et ce modèle est suffisamment significatif pour que la situation soit considérée comme préoccupante [36-39]

Dans le cas de la détection de situations anormales préalablement identifiées, on définit un modèle pour chacune de ces situations considérées comme critiques pour le patient, les autres étant alors considérées comme normales par défaut. Cela suppose d'intégrer dans le système la connaissance de toutes ces situations identifiées comme dangereuses. Cette méthode est fiable pour la détection des situations critiques connues, ce qui signifie que le taux de fausses détections ou faux positifs (la situation est considérée à tort comme critique) est faible. Par contre, elle ne permet pas la détection des situations

critiques qui ne sont pas préalablement connues du système : le taux de non détection ou faux négatifs (la situation est critique mais identifiée comme normale) est de ce fait élevé.

La seconde méthode envisagée consiste à établir un modèle du comportement habituel de la personne à domicile : on définit ainsi en quelque sorte le profil du patient. Chaque situation courante est identifiée par rapport à ce profil. Les situations considérées comme critiques sont celles correspondant à un écart significatif par rapport au modèle. Les situations critiques n'ont donc plus besoin d'être prédéfinies : le système est capable de détecter tout type de situation inhabituelle. On diminue ainsi le taux de situations critiques non détectées ou faux négatifs par rapport à la méthode précédemment citée. Par contre, la complexité de la modélisation du comportement habituel du patient induit le risque de mal identifier des situations pourtant normales, ce qui augmente le taux de fausses alarmes ou faux positifs.

Pour la détection des situations critiques d'une personne à domicile, il paraît important de veiller à réduire au minimum le taux de non détections (faux négatifs) afin de prendre le moins de risques possible quant à la sécurité du patient télésurveillé. Par ailleurs, l'acquisition de données relatives au comportement habituel du patient est tout à fait envisageable alors qu'il ne semble pas évident de définir exhaustivement toutes les situations critiques possibles. Il apparaît finalement plus naturel de chercher à modéliser les évolutions habituelles de l'état du patient en considérant alors que tout ce qui n'est pas identifié comme habituel est préoccupant, voire critique, plutôt que d'essayer d'identifier toutes les situations critiques possibles et de considérer alors par défaut toutes les autres comme normales.

Il s'agit donc de construire un modèle représentatif du comportement habituel, c'est-à-dire normal, du patient et de définir une méthode permettant de situer un état courant par rapport à ce modèle. Plus la situation observée est estimée comme étant loin du modèle, plus elle est critique. Le problème de décision consiste ensuite à déterminer sur cette distance au modèle les principes de déclenchement des différents types d'alarmes (seuils de décision par exemple).

3.4.3 Définition d'un ensemble fini d'états

Si on segmente les ensembles de valeurs prises par les variables d'état à valeurs dans un ensemble continu en un nombre fini d'intervalles, on n'a donc plus qu'à considérer des variables pouvant prendre chacune un nombre fini de valeurs. Chaque intervalle défini pour une variable représente une modalité de cette variable. A chaque variable d'état X_i , $0 < i \leq n$, on associe l'entier n_i représentant le nombre de ses modalités. Les n variables d'états sont les n composantes du vecteur d'état du patient. Il y a donc un nombre fini N de vecteurs d'état ou d'états possibles pour le patient. Ce nombre N des vecteurs d'état possibles est déterminé par la formule suivante :

$$N = \prod_{0 < i \leq n} n_i$$

Le modèle d'évolution des états possibles du système peut alors être représenté par un graphe fini de N états. Les états sont tous mutuellement exclusifs. Ce type de modélisation définit tous les états possibles du patient et les liens qui existent entre eux. Les états du patient sont les nœuds du graphe, reliés entre eux par des arcs orientés modélisant les transitions possibles ou relations d'accessibilité entre les différents états. A chaque arc se terminant sur un état donné est alors associée une valeur représentative de l'accessibilité de cet état à partir des autres états du graphe. On définit ainsi une chaîne de Markov [40] modélisant l'évolution habituelle de la situation du patient. Les nœuds du graphe sont les N vecteurs d'états possibles. Les paramètres de ce modèle d'évolution sont les probabilités initiales de chacun des états du graphe ainsi que la matrice de transition définissant les probabilités de transition entre les états.

Une autre solution, plutôt que de prendre en compte tous les états possibles dans la chaîne de Markov, pourrait être de considérer l'ensemble des états formés par les différents types de situation possibles : « situation habituelle », « situation préoccupante » et « situation critique », auxquels on ajoute un état absorbant représentant le décès du patient. Une représentation de la chaîne de Markov correspondante est présentée sur la Figure 5.

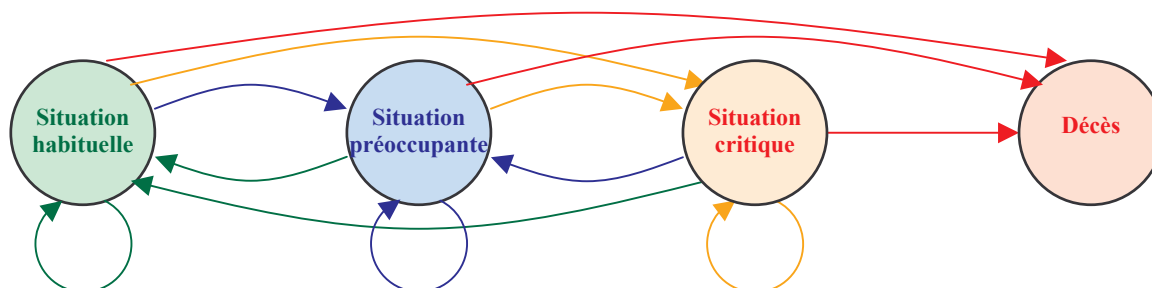


Figure 5. Chaîne de Markov représentant l'ensemble des situations possibles d'un patient

Il s'agit ensuite, à partir de l'observation d'un vecteur d'état X_t ou d'une suite de ces vecteurs de décider à quel type de situation (état caché de la chaîne de Markov sous-jacente) appartient la situation courante du patient. Cette vision du problème est liée à la théorie des chaînes de Markov cachées. Une de leur principale propriété est leur capacité à construire des modèles stochastiques sur des suites temporelles d'événements ou observations. Les types de situations possibles sont les états cachés du modèle. Les vecteurs d'états du patient sont les observations. Une succession d'observations ou vecteurs d'état permet d'évaluer la situation du patient par rapport à la chaîne de Markov sous-jacente, c'est-à-dire la chaîne de Markov cachée. Ce type de modélisation comporte trois types de paramètres à déterminer. Il faut tout d'abord être capable de définir les liens probabilistes qui unissent les différentes situations possibles du patient (états cachés) : c'est la matrice de transitions entre états. On doit également déterminer pour chaque état caché ou situation possible (« habituelle », « préoccupante » et « critique » par exemple) les probabilités d'observer chacun des vecteurs d'état possibles. Enfin, il faut déterminer les probabilités initiales de chacun des états cachés. Dans le cas présent, il semble normal de considérer que l'on est initialement presque sûrement dans la situation normale. Des algorithmes du type de celui de Viterbi [41] permettent ensuite de déterminer à partir d'une séquence d'observations (succession de vecteurs d'états) la suite d'états cachés la plus probable, ce qui induit la décision quant à la nature de la situation courante de la personne à domicile.

Cependant, dans le contexte de la télésurveillance médicale, on ne connaît bien que les évolutions habituelles du patient, sans trop d'idée des formes que peuvent présenter tout écart de comportement. L'utilisation d'une chaîne de Markov cachée composée des états « habituel », « préoccupant », « critique » et « décès » suppose également une bonne connaissance des situations préoccupantes et critiques puisqu'il faut être capable de quantifier les relations qui existent entre elles, ainsi que des conditions pouvant aboutir à la décès de la personne. Il faudrait pour cela disposer de nombreux scénarios de situations de patient évoluant vers ces différents états, ce qui n'est pas du tout le cas actuellement. Même si l'idée de situer une personne parmi un ensemble d'états correspondant aux types de situations que l'on cherche à identifier est attrayante, le manque de connaissances en ce qui concerne les évolutions critiques possibles incite à s'orienter plutôt vers la construction d'une chaîne de Markov complète composée des différents vecteurs d'état possibles et dont les relations d'accessibilité traduisent le comportement habituel du patient.

Le lien entre les deux types de modèle est peut-être simplement réalisé par un mécanisme d'abstraction de la chaîne de Markov la plus détaillée, c'est-à-dire celle où on choisit de représenter tous les états possibles d'un patient. Le principe de l'abstraction consiste à opérer des regroupements entre des états pour définir une nouvelle chaîne de plus petite taille. Les probabilités de transition du modèle de Markov ainsi obtenu sont calculées à partir des probabilités de transition du premier modèle. Un mécanisme d'abstraction ne se justifie cependant que dans le cas où il n'entraîne aucune perte d'information utile au problème de détection. Cela signifie qu'on observe les mêmes réponses du système avec les deux modèles, complet et simplifié.

Quel que soit le choix effectué pour la définition des états du modèle, les paramètres sont déterminés par apprentissage à partir de séquences de valeurs prises par les différentes variables d'état et représentatives d'une évolution habituelle de la personne. D'autres séquences de valeurs correspondant non seulement à des évolutions habituelles mais aussi à des scénarios d'évolution vers des situations critiques sont ensuite utilisées pour vérifier la validité du modèle ainsi défini. Il s'agit alors d'évaluer les performances de la décision pour savoir si elles sont satisfaisantes vis à vis des finalités du système. On considère notamment les taux de fausses détection (faux positifs) et de non détections (faux négatifs) (cf. Annexe G).

3.4.4 Le modèle de Markov

La chaîne de Markov est la théorie mathématique retenue pour être à la base de la détection des situations critiques de la personne à domicile. Le graphe fini des états de la personne à domicile est une représentation probabiliste des connaissances acquises sur l'évolution habituelle de l'état du patient.

Soit un système pouvant se trouver dans un nombre dénombrable d'états : $X = \{x_0, x_1, \dots, x_{N-1}\}$.

Entre les instants t et $(t+1)$, le système passe aléatoirement de l'état x_i à l'état x_j . Pour représenter ce processus aléatoire, on définit la variable X_t de la façon suivante : $X_t = x_i$ signifie que le système est dans l'état x_i à l'instant t .

Par définition, l'ensemble $Q = \{X_0, X_1, \dots, X_t, \dots\}$ est une suite stochastique à ensemble discret d'états [42].

Une chaîne de Markov est une suite stochastique à ensemble discret d'état satisfaisant à la propriété de Markov, c'est-à-dire telle que, pour une suite du premier ordre :

$$\Pr(X_t = x_j / X_{t-1} = x_i, X_{t-2} = \dots) = \Pr(X_t = x_j / X_{t-1} = x_i)$$

De même pour une suite du second ordre :

$$\Pr(X_t = x_k / X_{t-1} = x_j, X_{t-2} = x_i, X_{t-3} = \dots) = \Pr(X_t = x_k / X_{t-1} = x_j, X_{t-2} = x_i)$$

Une chaîne est dite homogène ou stationnaire si les probabilités de transition ne dépendent pas du temps. On peut alors poser :

$$p_{ij} = \Pr(X_t = x_j / X_{t-1} = x_i) \quad \text{pour l'ordre 1}$$

ou $p_{ijk} = \Pr(X_t = x_k / X_{t-1} = x_j, X_{t-2} = x_i)$ pour l'ordre 2

Pour résoudre la problématique de détection des situations critiques de la personne à domicile, on considère une chaîne de Markov stationnaire d'ordre 1. La prise en compte de la caractéristique dynamique du système (évolution des paramètres du modèle avec l'âge du patient par exemple) ne permet plus de considérer cette hypothèse de stationnarité comme exacte. Si on souhaite pouvoir néanmoins la conserver, il convient alors d'intégrer la composante temporelle à un autre niveau dans le processus de détection (au niveau du codage des variables d'état par exemple). Dans la première phase expérimentale, on ne considère pas les évolutions temporelles des paramètres du modèle.

La topologie d'une chaîne de Markov représente les transitions possibles entre les différents états. Il en existe différents types qui induisent des particularités au niveau de la matrice de transition entre les états. La propriété du modèle Gauche Droite par exemple est que l'état dans lequel se trouve le processus à modéliser augmente ou reste le même au cours du temps. Dans le cas de la modélisation de l'évolution d'un patient à domicile, on considère un modèle ergodique : ce modèle à la propriété que chaque état peut être atteint à partir de n'importe quel état en un nombre apériodique d'étapes. Une chaîne de Markov est apériodique si et seulement si il existe un instant \bar{n} et un état s tels que $p_{is}(\bar{n}) > 0 \quad \forall \text{ état } i$.

Les paramètres du modèle de Markov sont les probabilités initiales de chacun des états et les probabilités de transitions entre états. Le type de modèle identifié (modèle stationnaire et ergodique d'ordre 1) admet deux hypothèses :

- La distribution de probabilité d'un état à un instant $t+1$ ne dépend que de l'état à l'instant t . Les probabilités de transition entre états ne sont donc définies que par rapport à deux états successifs, sans tenir compte des évolutions aux instants antérieurs à t .
- La probabilité de transition d'un état E_i vers un état E_j est indépendante du temps.

En considérant une chaîne de Markov composée d'un ensemble fini de N états possibles E_j , $0 \leq j < N$, on définit alors les notions suivantes :

- $\forall i, 0 \leq i < N$, q_i est la probabilité initiale de l'état E_i . La distribution des probabilités initiales des états du modèle est donc $Q = | q_0 \quad q_1 \quad \dots \quad q_{N-1} |$. On a la propriété :

$$\sum_{i=0}^{N-1} q_i = 1$$

- $\forall i, 0 \leq i < N$, et $\forall j, 0 \leq j < N$, p_{ij} est la probabilité de transition de l'état E_i vers l'état E_j . Les probabilités de transition se présentent sous la forme d'une matrice de transitions de dimensions $N \times N$:

$$P = \begin{pmatrix} p_{00} & p_{01} & \dots & p_{0N-1} \\ p_{10} & p_{11} & \dots & p_{1N-1} \\ \vdots & \vdots & \dots & \vdots \\ p_{N-10} & p_{N-11} & \dots & p_{N-1N-1} \end{pmatrix}$$

On a la propriété :

$$\forall i, 0 < i \leq N, \sum_{j=1}^N p_{ij} = 1$$

3.4.5 Apprentissage des paramètres

L'apprentissage des paramètres du modèle est réalisé à partir de données de test. On considère par exemple une large fenêtre d'observation de l'état du patient composée d'une succession de M vecteurs d'état correspondant à une évolution normale de la personne à domicile : $\{\mathbf{X}_0, \mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_{M-1}\}$.

Les paramètres du modèle sont alors par exemple initialement déterminés par une méthode de comptage :

$$p_{ij} = \frac{M_{ij}}{M_i} \quad \text{et} \quad q_i = \frac{M_i}{M}$$

où M_{ij} est le nombre de couples de vecteurs d'état successifs \mathbf{X}_t et \mathbf{X}_{t+1} pour lesquels \mathbf{X}_t correspond à l'état E_i et \mathbf{X}_{t+1} à l'état E_j .

M_i est le nombre de couples de vecteurs d'état successifs \mathbf{X}_t et \mathbf{X}_{t+1} pour lesquels \mathbf{X}_t correspond à l'état E_i et \mathbf{X}_{t+1} à n'importe lequel des états du graphe.

M_i est le nombre de vecteurs d'état correspondant à l'état E_i .

M est le nombre total de vecteurs d'états considérés, c'est-à-dire la largeur de la fenêtre d'observation contenant les données d'apprentissage.

Une méthode basée sur la comparaison entre les états a également été proposée par Cao & Leong pour l'apprentissage des paramètres d'un modèle de Markov [14]. Cette méthode définit le calcul de probabilités de transition « de haut niveau » regroupant plusieurs transitions ayant le même type d'effet sur l'état d'un patient (amélioration, dégradation ou stabilité de l'état de la personne). Ce niveau

d'abstraction (*transition generalization*) est considéré comme nécessaire pour l'estimation des paramètres du modèle.

On pourra ensuite envisager le recours à des méthodes d'ajustement de ces paramètres à l'aide de données d'apprentissage supplémentaires.

3.4.6 Principe de numérotation des états du modèle

Pour modéliser l'état du patient, on construit une chaîne de Markov composée d'un nombre fini N d'états. On considère n variables d'état, chaque variable d'état X_i étant à valeur dans un ensemble discret de p_i éléments. Dans un premier temps, on fait l'hypothèse que toutes les variables sont codées en un même nombre d'éléments possibles p :

$$\forall i, p_i = p$$

Les intervalles de variation des variables d'état à valeur dans un ensemble continu sont donc partitionnés en p intervalles. La chaîne de Markov correspondante est alors composée de $N = p^n$ états représentant tous les vecteurs d'état possibles du modèle.

Pour la définition des paramètres de la chaîne de Markov (vecteur des probabilités initiales de chacun des états et matrice de transition entre états), il convient d'associer un numéro num_x , $0 \leq i < N$, à chaque vecteur d'état possible (ou état du modèle) $\mathbf{X} = [X_0 \ X_1 \ \dots \ X_{n-1}]$. On fait correspondre à chacun des p éléments de l'ensemble des valeurs possibles des variables d'état un entier k , $0 \leq k < p$, tel que :

si x_k est une valeur possible d'une variable d'état X_i , $0 \leq i < n$,

alors $x_k \in N$ et $0 \leq x_k < p$

on note $x_k = k$, $0 \leq k < p$

On peut alors considérer chaque vecteur d'état $\mathbf{x} = [x_0 \ x_1 \ \dots \ x_{n-1}]$ comme un nombre entier à n chiffres en base p , ω_x , obtenu par concaténation des valeurs de chaque variable d'état :

si $\mathbf{x} = | x_0 \ x_1 \ \dots \ x_i \ \dots \ x_{n-1} |$

alors $\omega_x = x_0 x_1 \dots x_i \dots x_{n-1}$ (base p)

L'ensemble des vecteurs d'état possibles décrivent alors tous les nombres à n chiffres (0 inclus) de la base p . Pour associer un numéro, et donc un nombre en base 10, à un vecteur d'état \mathbf{x} donné, il suffit alors de convertir le nombre ω_x (base p) en base 10. Ainsi, on obtient :

$$num_x = \sum_{i=0}^{n-1} x_i p^{n-i-1} \text{ avec } 0 \leq x_i < p$$

On considère par exemple 3 variables d'état ($n = 3$) toutes segmentées en 3 intervalles ($p = 3$) : « habituel » ($k = 0$), « diminué » ($k = 1$) et « augmenté » ($k = 2$). Le numéro num_x correspondant au vecteur d'état $\mathbf{x} = [\text{« } X_0 \text{ diminué » } \ \text{« } X_1 \text{ habituel » } \ \text{« } X_2 \text{ augmenté » }] = [1 \ 0 \ 2]$ est alors :

$$num_x = 1 \times 3^2 + 0 \times 3^1 + 2 \times 3^0 = 8$$

3.5 Identification de la situation du patient par rapport au modèle d'évolution

L'identification de la situation de la personne à domicile est réalisée à partir du vecteur d'état du patient à un instant donné et aux instant précédents et en référence à un modèle d'évolution habituelle. L'objectif est de déterminer si la situation observée est normale ou non pour le patient.

3.5.1 Description générale de la méthode

La situation du patient à un instant donné est représentée par la donnée d'une succession de vecteurs d'état définissant une fenêtre d'observation de durée T : $\{\mathbf{X}_{t-T+1}, \dots, \mathbf{X}_{t-2}, \mathbf{X}_{t-1}, \mathbf{X}_t\}$.

L'état de croyance du système, c'est-à-dire l'évaluation de la situation courante par rapport au modèle d'évolution, est défini à partir de cette fenêtre d'observation. L'état de croyance correspond à la probabilité d'occurrence de la fenêtre d'observation considérée en référence au modèle. On considère une fenêtre de durée T pour prendre une décision. Chaque nouvel instant $(t+1)$ entraîne la définition d'un nouveau vecteur d'état \mathbf{X}_t puis la révision de l'état de croyance du système. La fenêtre d'observation est décalée d'une unité pour prendre en compte le nouveau vecteur d'état et la situation du patient est réévaluée.

A l'initialisation du processus de détection, les vecteurs d'état utilisés pour l'apprentissage du modèle d'évolution sont inclus au début de la fenêtre d'observation. Cette fenêtre est ensuite décalée d'une unité à chaque nouveau vecteur d'état pris en compte.

On met ainsi en évidence l'existence de deux paramètres importants pour l'identification d'une situation : la définition des instants t de prise en compte d'un nouveau vecteur d'état (ce sont les instants de décision) et la détermination de la meilleure taille de fenêtre d'observation, c'est-à-dire du nombre de vecteurs d'états successifs à prendre en compte pour l'identification d'une situation à un instant donné.

3.5.2 Précision et Fiabilité

Une identification précise de la situation du patient consiste à décider directement à partir de sa distance au modèle si la situation correspond certainement à l'une des situations de référence : habituelle ou critique par exemple. Les probabilités d'occurrence de chaque situation de référence sont alors toutes nulles sauf pour la situation précisément identifiée pour laquelle cette probabilité vaut 1.

Plus on cherche à obtenir une identification précise de la situation du patient par rapport au modèle d'évolution, plus la décision quant à sa nature est facile à prendre, mais la fiabilité de cette décision diminue. Les probabilités d'erreur d'identification et de décision, c'est-à-dire les risques de fausse détection (faux positifs) et de non détection (faux négatifs), sont alors importantes. Une identification plus fiable est aussi en même temps beaucoup plus imprécise. Si par exemple on accorde la même probabilité d'occurrence à chaque type de situation, on a moins de risque de se tromper mais l'identification de la situation est en même temps très imprécise. Il y a donc un compromis à trouver entre précision et fiabilité.

L'entropie de Shannon, S_p , est une grandeur qui quantifie l'imprécision d'une mesure de probabilité sur un ensemble d'hypothèses d'état mutuellement exclusives $\{H_i\}_{i \in I}$ que sont les hypothèses sur les situations de référence :

$$S_p = \sum_{i \in I} \Pr(H_i) \cdot \log_2 (\Pr(H_i))$$

Son maximum est obtenu pour la mesure uniforme :

$$\forall i \in I, \Pr(H_i) = \frac{1}{\text{card}(I)}$$

Le maximum de l'entropie de Shannon correspond à l'hypothèse la plus imprécise.

3.5.3 Mesure de distance au modèle

Dans la modélisation de l'évolution normale du patient par une chaîne de Markov, positionner la situation courante par rapport au modèle consiste à calculer la probabilité d'existence, dans le cas d'une évolution normale du patient, de la suite des états qui composent la fenêtre d'observation considérée. Cette mesure de probabilité représente l'adéquation de la situation observée avec le modèle d'évolution normale du patient.

Afin de situer une situation courante du patient par rapport au modèle de Markov, on considère une fenêtre d'observation composée d'une succession de T vecteurs d'état du patient : $\{\mathbf{X}_{t-T+1}, \dots, \mathbf{X}_{t-2}, \mathbf{X}_{t-1}, \mathbf{X}_t\}$.

On calcule alors la probabilité d'occurrence de cette succession d'états par rapport aux paramètres du modèle. Dans le cas où les états du modèle de Markov sont les vecteurs d'état possibles pour le patient, cette probabilité d'occurrence d'une succession de vecteurs d'état peut être traduite simplement par le produit des probabilités de transition entre les états observés [43] :

$$d_t = \Pr(\mathbf{X}_{t-T+1}, \dots, \mathbf{X}_{t-2}, \mathbf{X}_{t-1}, \mathbf{X}_t) = q_{x_{t-T+1}} \prod_{\tau=t-T+2}^t p_{x_{\tau-1}x_\tau}$$

La distance d'une fenêtre d'observation du patient au modèle d'évolution habituelle est ainsi représentée par une mesure de probabilité. On a donc la propriété suivante : $0 \leq d_t \leq 1$.

3.5.4 Adaptation des paramètres du modèle

Si on considère l'évolution habituelle de la personne à domicile, on remarque que les états les « plus anormaux » ne doivent jamais être atteints. Cela induit donc de prendre en compte, suite à l'apprentissage, des probabilités de transition nulles dans les paramètres du modèle de Markov. Si par exemple l'état E_j ne doit pas être atteint dans une situation normale, toutes les probabilités de transition vers cet état doivent être nulles :

$$\forall i, 0 \leq i < N, p_{ij} = 0$$

La $j^{\text{ième}}$ colonne de la matrice de transitions P est donc nulle.

Le calcul de la probabilité d'occurrence d'une succession d'états définissant la distance d'une observation au modèle d'évolution, $\Pr(\mathbf{X}_{t-T+1}, \dots, \mathbf{X}_{t-2}, \mathbf{X}_{t-1}, \mathbf{X}_t)$, est le résultat du produit des probabilités de transition entre les états successifs et de la probabilité initiale du premier état considéré :

$$d_t = \Pr(\mathbf{X}_{t-T+1}, \dots, \mathbf{X}_{t-2}, \mathbf{X}_{t-1}, \mathbf{X}_t) = q_{x_{t-T+1}} \prod_{\tau=t-T+2}^t p_{x_{\tau-1}x_\tau}$$

Si une transition *a priori* impossible entre deux états apparaît dans une fenêtre d'observation, la probabilité de transition correspondante est nulle :

$$\exists \tau, t-T+2 \leq \tau \leq t, \text{ tel que } p_{x_{\tau-1}x_\tau} = 0$$

La probabilité d'occurrence de cette succession d'états $\Pr(\mathbf{X}_{t-T+1}, \dots, \mathbf{X}_{t-2}, \mathbf{X}_{t-1}, \mathbf{X}_t)$ est alors forcément nulle quels que soit les autres états composant la fenêtre d'observation. Jusqu'à présent, aucun traitement sur les valeurs récoltées pour les différentes variables d'état n'est envisagé. Les problèmes déjà soulevés concernant l'incertitude liée aux données (cf. §3.2.1) ne sont donc pas pris en compte et il est ainsi possible de rencontrer des états ou transitions aberrantes dans une fenêtre d'observation du patient. L'occurrence d'une transition estimée impossible lors de l'apprentissage, et donc le calcul d'une distance au modèle égale à zéro pour les fenêtres d'observation qui la contiennent, peut ainsi être la conséquence d'une valeur aberrante dans une série de données.

Pour le moment, on cherche simplement à éviter cette succession possible de zéros dans le calcul des distances au modèle. On décide donc de remplacer les probabilités qui devraient être nulles par des probabilités presque nulles, c'est-à-dire très faibles, pour atténuer l'influence de ces éventuelles données aberrantes :

$$p_{x_{\tau-1}x_\tau} = 10^{-s}, \quad s = 3 \text{ ou } 5 \text{ par exemple}$$

Cette transformation induit une erreur dans la matrice de transitions puisque la propriété :

$$\forall i, 0 < i \leq N, \sum_{j=1}^N p_{ij} = 1$$

n'est alors plus vérifiée. On considère cependant cette erreur négligeable.

3.6 Décision sur l'état du patient

3.6.1 Description générale du processus de décision

L'état de croyance du système, c'est-à-dire l'évaluation de la situation courante par rapport au modèle d'évolution, permet de prendre une décision globale sur la nature de la situation observée. Si celle-ci est critique, elle entraîne alors l'émission d'une alarme au médecin traitant et/ou à un centre de soins.

L'identification de la situation courante par rapport au modèle d'évolution est donc la base de la décision prise quant à la nature de l'état du patient. Il s'agit de classer cette situation parmi un ensemble de situations mutuellement exclusives possibles ou situations de référence : « situation habituelle », « situation préoccupante », et « situation critique » par exemple. Les situations qui ne sont pas habituelles déclenchent une alarme différente selon le type de situation observée. On remarque que l'on peut définir une relation d'ordre sur l'ensemble des situations de référence considérées. Elles peuvent en effet être ordonnées par risque croissant pour le patient : « situation habituelle » puis « préoccupante » et enfin « critique » par exemple.

L'ensemble de décisions possibles, $\{D_j\}_{0 < j \leq q}$, sur la situation du patient à un instant donné correspond à l'ensemble des types de situation envisagés : c'est l'ensemble des situations de référence ou classes de situations possibles.

On définit ensuite l'ensemble des aléas $\{\omega_i\}_{0 < i \leq p}$ qui représentent l'incertitude du système de décision (imprécision). Les aléas correspondent aux mesures de probabilités sur les différentes hypothèses de situation du patient. Il s'agit de déterminer la probabilité pour une situation observée d'appartenir à chacun des types de situation envisagés : situation « habituelle », « préoccupante », ou « critique » par exemple. On détermine donc autant d'aléas que de classes de situation.

L'évaluation de la situation courante du patient permet de le situer par rapport à son modèle d'évolution habituelle. C'est cette évaluation qui permet de déterminer la valeur des aléas, c'est-à-dire la probabilité pour le patient d'être dans chacun des types de situation envisagés. Si on considère par exemple trois situations possibles {"habituelle", "préoccupante", et "critique"}, on définit trois aléas : ω_1 est la probabilité d'être dans une situation habituelle, ω_2 d'être dans une situation préoccupante et ω_3 d'être dans une situation critique.

Dans le cas d'une modélisation par une chaîne de Markov, on dispose par exemple d'une mesure de probabilité pour positionner la situation courante par rapport au modèle : c'est la probabilité pour une suite d'états donnée d'être en adéquation avec le modèle d'évolution habituelle du patient. Plusieurs démarches sont alors envisageables pour prendre une décision selon la précision que l'on décide d'accorder au rattachement d'une situation courante à l'ensemble des types de situation envisagés ou situations de référence : la détermination de seuils de décision ou l'utilisation d'arbres de décision.

3.6.2 Seuils de décision

La gravité d'une situation observée est proportionnelle à la distance entre cette situation et le modèle d'évolution habituelle du patient. La probabilité d'adéquation d'une situation au modèle, qui mesure cette distance au modèle, est donc d'autant plus faible que la situation observée est loin du modèle, et donc grave. On peut ainsi envisager de prendre une décision à partir de la détermination de seuils de décision sur cette probabilité représentant la proximité au modèle. On détermine une partition de l'intervalle $[0,1]$ des valeurs possibles de cette probabilité en un nombre d'intervalles disjoints correspondant au nombre de situations envisagées. On assigne alors à chaque intervalle une et une seule situation de référence. La valeur de la probabilité définissant la distance au modèle appartient forcément à l'un de ces intervalles, ce qui détermine précisément la classe de la situation observée, et donc la décision à prendre.

Par exemple si on considère un ensemble de décisions de cardinal égal à trois : {"situation habituelle", "situation préoccupante", "situation critique"}, il faut alors segmenter l'intervalle $[0,1]$ en trois intervalles disjoints et assigner à chacun d'entre eux l'une des situations envisagées. On peut par exemple décider que la situation observée est « habituelle » si la probabilité d'adéquation au modèle appartient à l'intervalle $]0.5,1]$, « préoccupante » pour $]0.25, 0.5]$ et enfin « critique » pour $[0, 0.25]$.

La définition d'une partition permet de déterminer avec précision la situation dans laquelle se trouve le patient. L'ensemble des aléas du système, c'est-à-dire l'ensemble des mesures de probabilité sur les situations de référence $\{\omega_i\}_{0 < i \leq p}$, ne comprend donc que des valeurs nulles sauf une égale à 1. D'après l'exemple précédent, si on a par exemple une probabilité d'adéquation au modèle égale à 0.42, on estime que la situation est préoccupante, et seule la valeur de l'aléa ω_2 est alors non nulle et égale à 1. La prise de décision D_j correspond à l'aléa ω_i non nul :

$$D_j \text{ est tel que } j = i, \text{ avec } \omega_i = 1$$

Si $\omega_2 = 1$ et $\omega_i = 0 \forall i \neq 2$, on décide donc de prendre la décision D_2 : « la situation est préoccupante ».

La détermination et l'ajustement des seuils de décision est effectuée à l'aide de données d'apprentissage dont on connaît la classe d'appartenance *a priori*, c'est-à-dire la situation de référence : « habituelle », « préoccupante » ou « critique » par exemple. Cela revient à considérer un ensemble de fenêtres d'observations reflétant différents types d'évolution du patient. Chacune de ces fenêtres est donc rattachée à une situation de référence. On calcule alors la distance de chacun de ces ensembles de vecteurs d'état au modèle d'évolution normale du patient. L'idée est ensuite d'essayer de discriminer parmi l'ensemble des distances ainsi obtenues des groupes de valeurs qui correspondraient à chaque situation de référence. La détermination numérique des frontières entre ces groupes correspond alors à la définition des seuils de décision.

3.6.3 Arbres de décision

La détermination de partitions de l'intervalle de valeurs de la probabilité d'adéquation d'une situation observée au modèle d'évolution normale peut cependant paraître un peu brutale. A partir de cette probabilité et peut-être d'autres critères liés au contexte de la décision (paramètres environnementaux par exemple), on peut aussi à partir de chaque état observé attribuer un coût à chaque décision, c'est-à-dire à chaque type de situation envisagé (situations de référence). L'ensemble des aléas du système, $\{\omega_i\}_{0 < i \leq p}$, ne comprend alors que des valeurs *a priori* non nulles. Chaque aléa ω_i représente le degré d'appartenance de la situation courante au type de situation correspondant. L'aléa ω_2 est par exemple le degré d'appartenance de la situation courante au type de situation « préoccupante ».

La prise d'une décision D_j sur l'aléa ω_i peut alors être quantifiée en terme d'utilité ou de conséquence possible de la décision. Cela correspond à envisager le cas où on décide D_j quand ω_i est réalisé. Cette valeur numérique peut par exemple représenter un gain, un risque ou une perte. On forme ainsi une matrice des utilités ou des conséquences des décisions possibles pour un état donné : $\{U_{ij}\}_{0 < i \leq p, 0 < j \leq q}$.

Le problème de décision est alors représenté par un arbre de décision (Figure 6) construit à partir d'une situation courante et du modèle d'évolution du système. L'ensemble des hypothèses d'état prises en compte à l'instant de décision (aléas), ainsi que l'ensemble des décisions possibles varient au cours du temps.

Les nœuds ronds de l'arbre (\circ) représentent les éventualités sur l'état du patient ($\{\omega_i\}$), c'est-à-dire les degrés d'appartenance de la situation observée à chaque situation de référence. Les nœuds suivants, carrés (\square), correspondent à un choix du système de décision à partir de ces éventualités ($\{D_j\}$).

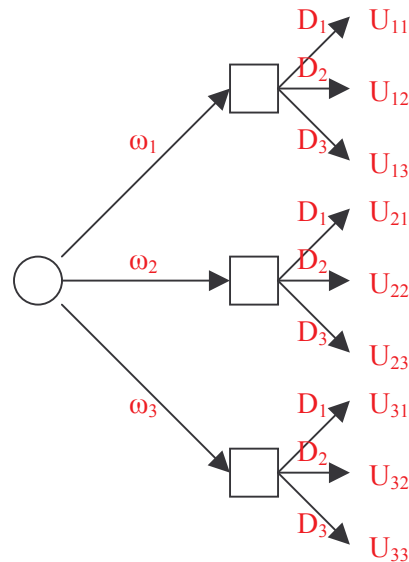


Figure 6. Ensemble des conséquences possibles U_{ij} des décisions D_j sur les aléas ω_i

Les stratégies définissent ensuite les règles du choix d'une décision à chaque étape, c'est-à-dire à chaque instant de décision. Une fois quantifiées l'utilité ou les conséquences possibles de chaque décision sur l'ensemble des aléas, on applique un ou plusieurs critères de décision les plus adaptés à la situation (minimisation des risques, stratégie la plus prudente par exemple). Quelques critères couramment utilisés dans les problèmes de décision sont la maximisation de l'espérance mathématique, le maximum de vraisemblance, les critères du Min-Max et du Max-Min.

Le choix des critères de décision dépend des caractéristiques et des besoins de l'application. Une règle de décision est évaluée notamment par les taux de fausses détections (faux positifs) et de non détections (faux négatifs). Ces taux sont plus ou moins critiques selon le contexte.

3.7 Prise en compte de la composante temporelle

3.7.1 Définition des instants de décision

La situation du patient évolue de façon continue au cours du temps. Ces évolutions sont connues du système à partir des nouvelles informations délivrées en sortie des capteurs (informations perceptives) ou bien par notification de modifications effectuées sur les données cliniques du patient par un des acteurs du système SIC-HIS (informations cliniques). Il faut donc réviser en permanence les hypothèses d'état ou état de croyance pour intégrer les nouvelles informations délivrées au système.

Le vecteur $\mathbf{X}_t = [X_{t,1} X_{t,2} \dots X_{t,n}]$ est la représentation mathématique significative de l'état du patient à un instant t donné. Ses composantes sont les variables d'état du patient. Elles sont fonction des observables, c'est-à-dire des données issues des capteurs et des données cliniques du patient : $\{Y_{t,i}\}_{0 < i \leq m}$. L'intégration de nouvelles données à un instant suivant ($t+1$) engendre une nouvelle famille de données $\{Y_{t+1,i}\}_{0 < i \leq m}$. Ces données induisent la prise en compte du nouveau vecteur d'état \mathbf{X}_{t+1} .

La situation du patient à un instant t est représentée par la donnée d'une succession de vecteurs d'état définissant une fenêtre d'observation de durée T : $\{\mathbf{X}_t, \mathbf{X}_{t-1}, \mathbf{X}_{t-2} \dots \mathbf{X}_{t-T+1}\}$. C'est à partir de ces données qu'est évaluée la situation du patient, évaluation suivie d'une prise de décision sur la nature de cette situation. Une question importante est de savoir à quels instants il est nécessaire de réévaluer la situation du patient, et donc de prendre éventuellement une nouvelle décision.

On peut penser que chaque nouvelle donnée reçue ou modifiée dans la base de données du système HIS doit induire une révision des croyances quant à la situation de la personne à domicile. Pour ce qui concerne les données cliniques, cette démarche semble effectivement pouvoir s'appliquer. Les données du dossier médical du patient utilisées dans le processus de détection de situations critiques sont peu

nombreuses et *a priori* rarement modifiées. La modification de la valeur d'une de ces données peut donc s'avérer stratégique et il est alors nécessaire de réévaluer rapidement la situation du patient.

Par contre, les données obtenues via les capteurs sont délivrées au système en permanence et de manière asynchrone. La révision de la décision à chaque nouvelle donnée reçue risque d'entraîner une surcharge du processus de décision et de ce fait une diminution des performances du système. La réception de certains types de données peut cependant être stratégique. Considérant le cas de réception d'un indicateur de chute et d'arrêt cardiaque, il faut alors prendre en compte immédiatement cette nouvelle donnée et d'ailleurs également déclencher immédiatement une alarme. Cet exemple met en évidence deux situations :

- Certaines données reçues des capteurs comme l'indication d'une chute doivent entraîner une réévaluation immédiate de la situation de la personne à domicile, entraînant éventuellement une nouvelle prise de décision et peut-être le déclenchement d'une alarme.
- D'autres données doivent même court-circuiter le processus de décision pour déclencher directement une alarme : c'est le cas d'un arrêt cardiaque par exemple.

Cette remarque met de nouveau en évidence la distinction entre deux processus de détection parallèles : la détection des évolutions critiques du patient et celle des événements considérés directement comme alarmants et qui déclenchent automatiquement une alarme. Une valeur de variable d'état qui atteint une zone d'urgence donne lieu à l'envoi immédiat d'une alarme : c'est le cas par exemple d'une crise hypertensive correspondant à une valeur supérieure à 220 pour la pression systolique ou supérieur à 130 pour la pression diastolique (cf. Annexe C.3.1).

Indépendamment du cas particuliers de la réception de données déclenchant automatiquement une alarme, il faut déterminer des intervalles de décision définissant les instants de déroulement du processus de détection des situations critiques de la personne à domicile.

La méthode la plus simple est de considérer des intervalles de décision de taille fixe. Seule la réception d'une donnée alarmante peut alors entraîner une décision avant l'échéance de la durée fixée entre deux décisions. Cette situation n'est cependant pas très satisfaisante pour plusieurs raisons liées au fait que le risque encouru par la personne à domicile n'est pas toujours le même. Le risque de chuter par exemple augmente avec l'activité du patient. Si la personne est en train de dormir, le risque de chuter est nul alors que si elle fait le ménage de son appartement, ce risque est fortement accru. Le rythme de vie de la personne à domicile est donc fortement lié à l'éventualité de détection d'une situation inhabituelle. On peut alors imaginer définir la taille des intervalles de décision comme une fonction du temps afin par exemple de ne pas déclencher le processus de décision pendant la nuit, sauf si une donnée alarmante est reçue. On définit ainsi des périodes sensibles pendant lesquelles on accroît la surveillance du patient.

Une autre remarque est que le risque pour un patient de faire face à une situation critique, voire urgente, augmente avec la gravité des situations détectées aux instants de décision précédents. Tant que la situation de la personne à domicile est normale, le risque d'une évolution anormale est moindre que si cette situation avait déjà été évaluée comme préoccupante par exemple. Cela suppose donc que plus la gravité de la situation évaluée à un instant donné augmente, plus les instants de décision doivent être rapprochés. La taille des intervalles de décision doit donc diminuer quand la gravité de la situation du patient augmente.

3.7.2 Limite de validité du modèle d'évolution

On peut également se poser la question de la validité du modèle d'évolution normale du patient à long terme. Les habitudes comportementales d'une personne évoluent en permanence. En vieillissant par exemple, une personne est amenée à moins se déplacer, peut-être à prendre du poids ou à avoir un rythme cardiaque plus élevé. Ce sont des évolutions lentes et pas vraiment inquiétantes qui doivent remettre en question la délimitation de la région considérée comme « normale » pour cette personne, c'est à dire le modèle représentatif de son évolution habituelle.

Cela signifie qu'il convient « à un moment donné » non seulement de réviser les intervalles de segmentation des variables continues définissant par exemple les intervalles de poids « diminué »,

« habituel » et « augmenté », mais également de rejouer le processus d'apprentissage permettant de déterminer les paramètres du modèle. On peut aussi envisager de réviser la règle de décision quant à la nature de la situation observée. De nouvelles données d'apprentissage sont nécessaires à ces tâches de révision. Le plus simple semble de considérer les vecteurs d'états observés aux instants précédents comme les nouvelles données d'apprentissage. Cela suppose qu'on est certain, dans une certaine marge de confiance, que ces données sont vraiment représentatives d'un comportement normal de la personne à domicile. Si la probabilité d'adéquation d'une situation avec le modèle d'évolution normale est proche de 1, c'est qu'on est presque certain que les T vecteurs d'état qui ont servi au calcul de cette probabilité, $\{X_t, X_{t-1}, X_{t-2}, \dots, X_{t-T+1}\}$, correspondent à une évolution habituelle du patient.

Il s'agit alors de déterminer cette marge de confiance à l'intérieur de laquelle on considère que les derniers vecteurs d'état observés sont suffisamment représentatifs du comportement normal du patient pour autoriser leur utilisation comme données d'apprentissage afin d'affiner les paramètres du modèle. Cette marge de confiance peut être représentée par un seuil correspondant à la valeur minimale tolérée pour la probabilité d'adéquation d'une situation au modèle d'évolution afin d'autoriser cette réévaluation. Ce seuil est distinct de celui délimitant les situations normales, défini dans le cas où on décide de déterminer des seuils de décision pour estimer la situation courante de la personne à domicile. On peut par exemple décider qu'une situation est normale si la probabilité d'adéquation d'une situation au modèle appartient à l'intervalle $]0.5, 1]$. Cependant, on peut estimer qu'une situation suffisamment normale pour permettre une réévaluation des paramètres du modèle doit elle correspondre à une probabilité d'adéquation appartenant à l'intervalle $]0.95, 1]$ par exemple. Il faut être plus exigeant sur la « normalité » d'une situation quand il s'agit d'utiliser ses composantes, les vecteurs d'état du patient, à des fins d'apprentissage.

Parallèlement à la prise en compte de nouvelles données d'apprentissage, il convient à l'inverse d'en oublier certaines qui ne sont plus suffisamment représentatives du patient compte-tenu de l'évolution des paramètres du modèle. On peut soit décider d'éliminer systématiquement les M premiers vecteurs d'apprentissage quand on en prend en compte M nouveaux, soit appliquer sur les premiers vecteurs d'apprentissage les mêmes critères de fiabilité que ceux utilisés pour la sélection de nouvelles données. Ils ne sont alors conservés dans la série de données d'apprentissage que si la distance au modèle d'évolution calculée à partir de ces vecteurs appartient à l'intervalle de confiance défini à l'intérieur de la zone de variation normale des distances au modèle : c'est l'intervalle $]0.95, 1]$ de l'exemple précédent, défini à l'intérieur de l'intervalle $]0.5, 1]$ représentant l'ensemble des distances habituelles d'un patient par rapport à son modèle d'évolution.

3.7.3 Interprétation des écarts comportementaux

Le modèle d'évolution de la personne à domicile est le reflet de son comportement habituel. Cette situation habituelle du patient n'est pas forcément normale au sens des définitions empiriques de la médecine. Un écart par rapport au modèle n'est alors pas toujours dangereux. Il peut en effet correspondre à un retour du patient à des valeurs « plus normales » pour certaines de ces variables. Dans l'exemple du patient hypertendu, ce genre de situation pourrait correspondre à une baisse de tension suite à un traitement relatif à l'hypertension par exemple. Il faut dans ce cas informer le médecin traitant et réévaluer simplement le modèle d'évolution afin de ne plus détecter ce genre d'écart comme correspondant à une situation préoccupante. Par contre, un écart au modèle qui éloigne les valeurs des variables d'état du patient des intervalles normaux de variation définis généralement par la profession médicale peut devenir vraiment critique.

La méthode de modélisation du comportement habituel d'un patient décrite dans les paragraphes précédents ne permet pas *a priori* de définir si un écart observé par rapport au modèle correspond à un retour à des valeurs normales ou non : cette décision est laissée au médecin traitant. Il dispose pour cela notamment de la visualisation des courbes de tendance de chaque variable d'état, ainsi que d'un ensemble de messages envoyés par le système à chaque instant (« hypertension artérielle modérée » ou « obésité légère » par exemple).

Pour automatiser cette démarche, on pourrait imaginer disposer d'un modèle d'évolution habituelle d'une personne correspondant exactement (ou « le mieux possible ») aux critères définis comme « normaux » par la médecine. Il faudrait ensuite pouvoir calculer la distance entre un modèle d'évolution habituelle d'une personne à domicile et ce modèle de référence. Si suite à un écart par rapport à son propre modèle

d'évolution, un patient se rapproche du modèle de référence, c'est que l'évolution de son comportement tend vers une évolution médicalement « plus normale ».

3.8 Evaluation des performances du système

Le processus de détection de situations critiques pour le patient est évalué pour mettre en évidence les avantages et inconvénients des différentes méthodes utilisées. Cela donne des indications sur les orientations à prendre pour essayer d'améliorer les performances du système.

Il faut donc définir un ensemble d'indicateurs de performance pertinents, c'est-à-dire significatifs par rapport aux finalités du système. Les mesures les plus pertinentes sont celles qui permettent de vérifier la fiabilité du système au niveau des points sensibles du problème considéré.

Des indicateurs efficaces à considérer sont les taux d'erreurs et de bonnes détections. Il y a deux types d'erreurs pouvant survenir : les fausses alarmes (ou faux positifs) et les non détections (ou faux négatifs) (cf. Annexe G). Pour la fiabilité du système de détection des situations à risque d'une personne à domicile, il est important que le taux de non détections soit faible pour prendre le moins de risque possible sur la santé du patient. Par ailleurs, le taux de fausses alarmes doit lui aussi être le plus faible possible pour la crédibilité de la décision. Un système qui répond souvent par des alarmes a tendance à être finalement ignoré du personnel médical qui ne lui accorde plus aucune crédibilité. Cependant, en cherchant à diminuer le taux de fausses alarmes on augmente par ailleurs le taux de non détection et inversement. Il convient donc de trouver un compromis entre les valeurs de ces deux taux en fonction de l'importance accordée à chacun d'entre eux dans le contexte du projet.

La méthode COR (Caractéristique Opérationnelle du Récepteur), appelée encore méthode *ROC (Receiver Operating Characteristic)*, de Provost & Fowcett [44] exploite les relations entre les taux de non détections et de fausses alarmes pour évaluer les performances d'un système de décision. Une courbe COR (cf. Figure 7) est une courbe paramétrique obtenue en faisant varier le seuil de décision et en calculant les probabilités de fausse alarme et de bonne détection à chaque instant. Le seuil de décision est le paramètre que l'on peut régler pour ajuster les taux de fausses alarmes et de non détections en fonction des contraintes de performance du système. La courbe COR pour un classificateur idéal est un rectangle, celle pour un classificateur aléatoire une droite sur la diagonale. L'intégrale de la surface sous la courbe COR est souvent utilisée comme mesure globale de la performance d'un algorithme de classification.

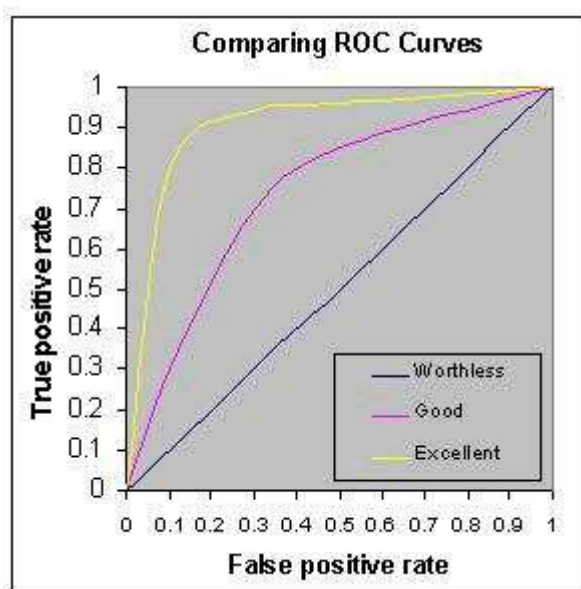


Figure 7. Exemple d'une courbe COR pour un mauvais, un bon et un excellent classificateur

Un autre critère peut-être important à considérer est la rapidité de détection d'une situation critique pour le patient.

3.9 Synthèse et schéma général du processus de décision

Le schéma de la Figure 8 résume les différentes phases du processus de synthèse de données en vue de la décision dans le cadre de la détection de situations critiques d'un patient. On y retrouve l'enchaînement des étapes précédemment citées :

- Intégration d'informations de différentes origines ;
- Représentation de l'état du patient ;
- Evaluation de la situation observée par rapport au modèle d'évolution habituelle ;
- Décision sur la nature de la situation observée dans l'objectif de détection des situations critiques.

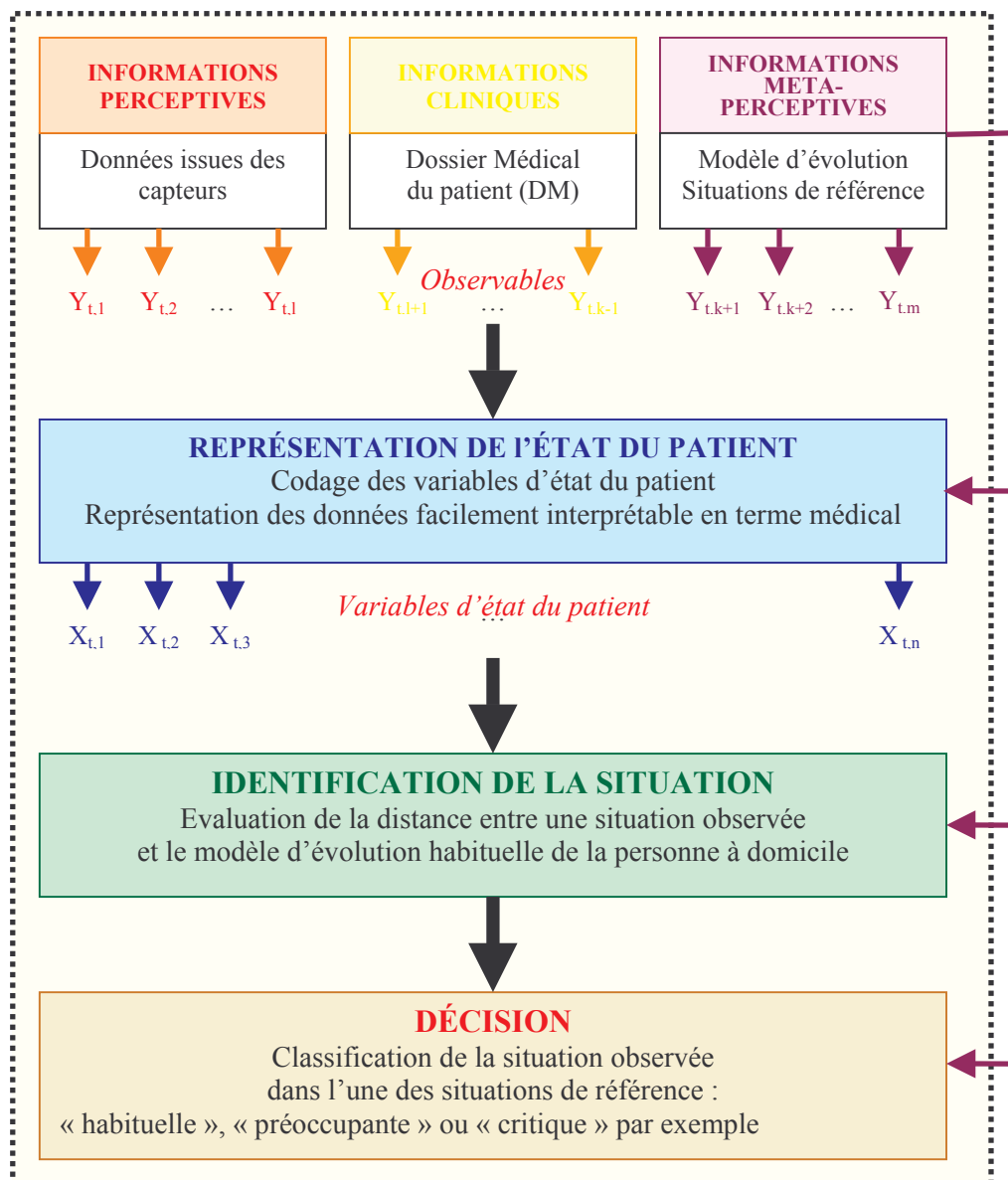


Figure 8. Les étapes du processus de synthèse de données et de décision pour la détection des situations critiques d'une personne à domicile

4 Apprentissage des paramètres du modèle

4.1 Paramètres stratégiques du système de décision

La description des grandes étapes du processus de synthèse de données a permis de cibler quelques domaines théoriques de recherche liés au problème de détection de situations à risques pour une personne à domicile. Le principe de détection d'une situation critique consiste à modéliser le comportement habituel d'un patient et à considérer un comportement suspect comme un écart par rapport à ce modèle.

Le processus de détection de situations critiques tel qu'il a été décrit dans les paragraphes précédents met en évidence un certain nombre de paramètres et principes fondamentaux ainsi que des critères de décision stratégiques pour les performances du système de décision :

- le nombre et la pertinence des variables d'état ;
- le nombre et les limites des intervalles de segmentation des variables d'état à valeurs dans un ensemble continu ;
- la méthode de modélisation de l'évolution habituelle du patient : dans le cas de l'utilisation d'une chaîne de Markov, les paramètres à déterminer sont les probabilités initiales des états et les probabilités de transition entre états ;
- les instants de décision ;
- la largeur de la fenêtre d'observation ;
- la mesure de distance entre une situation observée et le modèle ;
- le nombre et la nature des situations de référence considérées ;
- les principes de décision sur la distance au modèle ;
- les instants et les méthodes de révision du modèle (prise en compte de la composante temporelle).

Initialement, les valeurs de certains paramètres doivent être déterminées intuitivement, même si elles peuvent être modifiées ou affinées une fois les performances du système évaluées. C'est le cas par exemple de la définition des instants de décision, de la taille de la fenêtre d'observation et des situations de référence. D'autres paramètres sont déterminés directement par apprentissage à l'aide de séquences de données pertinentes : paramètres d'une chaîne de Markov, seuils de décision par exemple.

Les paragraphes suivants reprennent les étapes de la construction du modèle identifiées au paragraphe précédent (cf. §3), c'est-à-dire les étapes d'apprentissage des paramètres stratégiques du modèle :

- Définition des intervalles de segmentation pour la représentation de l'état du patient ;
- Définition des paramètres de la chaîne de Markov pour l'identification de la situation du patient ;
- Définition des principes de décision (seuils) sur la situation de la personne à domicile.

4.2 Données nécessaires à l'apprentissage et à la validation du modèle

La mise en œuvre du processus de détection, c'est-à-dire l'apprentissage du modèle, nécessite un grand nombre de données d'apprentissage et de validation. Elles sont de plusieurs types :

- Des séquences de vecteurs d'état représentatifs de l'évolution habituelle du patient (données d'apprentissage et de validation) ;
- Des séquences de valeurs correspondant à l'évolution de l'état du patient vers une situation critique (données de test).

Ces données sont utilisées à plusieurs niveaux lors de la mise en place du système. Elles sont notamment indispensables pour :

- La détermination des intervalles de segmentation des variables d'état à valeur dans un intervalle continu : il est par exemple nécessaire de disposer d'un ensemble de valeurs significatives d'un poids habituel pour le patient si on veut pouvoir délimiter un intervalle de valeurs considérées comme « habituelles ».
- La définition des paramètres du modèle d'évolution : détermination des probabilités des états initiaux et des probabilités de transition entre états.
- L'évaluation des principes de décision pour la classification d'une situation observée dans l'une des situations de référence.

Pour les deux premières démarches, on utilise les données d'apprentissage et de validation principalement. Les données de test sont par contre indispensables à la détermination des seuils de décision.

4.3 Description du processus d'apprentissage du modèle

Les paragraphes suivants décrivent les étapes du processus de définition et d'apprentissage du modèle pour la détection des situations critiques de la personne à domicile.

4.3.1 Définition des variables d'état du patient

Elle consiste en la définition d'un ensemble de variables d'état médicalement significatives et pertinente par rapport au problème d'évaluation de l'état du patient à domicile.

On décide de prendre en compte un ensemble de variables d'état correspondant à des mesures fréquemment utilisées par les médecins lorsqu'ils doivent juger l'état général d'une personne. Quelques uns de ces critères médicaux sont détaillés en annexes, comme les mesures cardiaques (Pressions Artérielles Diastolique – PD – et Systolique – PS –, Fréquence Cardiaque – FC –, cf. Annexe C), un indice de corpulence (Indice de Masse Corporelle – IMC –, cf. Annexe D), un indicateur de troubles de la miction (cf. Annexe E), etc.

On définit ainsi l'ensemble des variables d'état significatives en terme médical qui sont utilisées pour représenter à tout instant l'état du patient. Les valeurs de l'ensemble de ces variables d'état sont obtenues à partir des données issues des capteurs installés dans l'appartement équipé de la faculté de médecine de Grenoble (tensiomètre pour les pressions artérielles et la fréquence cardiaque, pèse personne pour le calcul de l'IMC, capteurs de présence pour détecter la pollakiurie, cf. Annexe B) et en utilisant les données cliniques du patient (taille pour le calcul de l'IMC par exemple) : ce sont les *observables*.

A ce type de variables s'ajoutent des paramètres environnementaux dont les valeurs sont obtenues à partir de la combinaison de données issues des capteurs : distance ou vitesse de déplacement de la personne, température ambiante par exemple.

4.3.2 Codage des variables d'état

Il s'agit de coder les valeurs prises par les variables d'état du patient afin de les qualifier de manière globale et significative en terme médical. Cette étape consiste donc à segmenter les intervalles de valeurs des variables d'état en intervalles de valeurs d'interprétation en terme médical. Cette segmentation est réalisée en adéquation avec la finalité du système qui consiste à détecter les situations inhabituelles du patient à domicile. On pense ainsi à diviser les valeurs prises par les variables continues médicalement significatives, comme le poids, en trois intervalles correspondant aux valeurs « habituelles », « diminuées » et « augmentées » par exemple. Plus simplement, on peut aussi envisager de segmenter cet intervalle en seulement deux parties : les valeurs « habituelles » et les autres « inhabituelles ». Cette dernière méthode ne permet cependant pas de différencier un certain nombre d'états et de transitions ayant pourtant une interprétation tout à fait différente pour le patient. Si on considère par exemple le cas d'un sujet hypertendu, les scénarios d'évolution correspondant d'une part à un succès thérapeutique suite à un traitement contre l'hypertension et d'autre part à une aggravation de sa situation sont alors tous deux considérés comme un même écart au modèle puisqu'ils correspondent à un passage à des valeurs « inhabituelles ». Ces deux scénarios d'évolution ne peuvent dès lors plus être différenciés. On pressent

donc que cette simplification n'est pas valide car elle entraîne une perte d'information utile vis à vis de la problématique de détection de situations critiques.

Deux paramètres critiques à déterminer sont ainsi soulignés :

- **La détermination du nombre d'intervalles de segmentation** : ce nombre induit directement le nombre d'états à prendre en compte pour la modélisation du comportement habituel du patient par un graphe fini d'états.
- **La délimitation de ces intervalles de valeurs** : elle nécessite de disposer d'un certain nombre de valeurs représentatives d'une situation habituelle pour le patient.

On a également remarqué la nécessité de réviser les intervalles de valeurs ainsi définis pour prendre en compte les évolutions naturelles des valeurs normales prises par les différentes variables, par exemple l'augmentation de la fréquence cardiaque avec l'âge.

4.3.3 Représentation de l'état du patient

Le codage des ensembles de valeur des variables d'état du patient permet de définir un nombre fini d'états, c'est-à-dire de vecteurs d'états possibles $\mathbf{X}_t = [x_0 \ x_1 \ \dots \ x_{n-1}]$, pouvant être atteints par la personne à domicile. On définit ainsi le nombre N d'états représentés dans la chaîne de Markov sous-jacente au problème de décision : il est déterminé par le nombre de modalités n_i de chaque variable d'état X_i après segmentation de l'intervalle de variation de ses valeurs (codage) :

$$N = \prod_{0 < i \leq n} n_i$$

Il s'agit ensuite d'associer à chaque vecteur d'état pouvant être atteint par le patient le numéro d'un état de la chaîne de Markov (cf. §3.4.6). Les ensembles de données d'apprentissage, de validation et de test sont ainsi codés en une succession de numéros d'états de la chaîne de Markov.

4.3.4 Définition des paramètres de la chaîne de Markov

L'évolution habituelle de la personne à domicile est modélisée par une chaîne de Markov. Ce modèle est un graphe fini d'états reliés entre eux par des arcs auxquels on associe des probabilités de transition.

On considère donc tous les états possibles pour le patient et on détermine les relations d'accessibilité entre ces états. Les paramètres d'une chaîne de Markov sont d'une part les probabilités initiales de chacun des états, et d'autre part les probabilités de transition entre états. Ces paramètres doivent être déterminés par apprentissage à partir d'une large fenêtre d'observation de l'état du patient composée d'une succession de M vecteurs d'état correspondant à une évolution habituelle de la personne à domicile : $\{\mathbf{X}_0, \mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_{M-1}\}$. Ils sont ensuite validés par d'autres séries de données du même type représentant des évolutions habituelles mais aussi critiques de la personne à domicile.

Par ailleurs, il a également été mis en évidence la nécessité de réviser les paramètres du modèle pour prendre en compte les évolutions naturelles du comportement habituel du patient. Il s'agit d'intégrer de façon explicite la composante temporelle dans le modèle.

4.3.5 Identification d'une situation par rapport au modèle

La situation du patient à un instant t est évaluée en terme de distance par rapport au modèle de son évolution habituelle et à partir d'une fenêtre d'observation du patient, c'est à dire une succession de vecteurs d'états codés en terme de numéro d'état en référence à la chaîne de Markov sous-jacente au problème de décision.

Plusieurs paramètres sont stratégiques pour cette étape du processus :

- **Le choix de la taille de la fenêtre d'observation**, qui correspond au nombre d'états du patient à prendre en compte à chaque instant de décision : $\{\mathbf{X}_0, \mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_{T-1}\}$ pour une fenêtre de largeur T . Ce choix est dépendant notamment de la durée définie entre deux instants de décision.

- La détermination d'une mesure de distance entre une séquence de vecteurs d'états (fenêtre d'observation) et le modèle d'évolution : cette distance est estimée par une mesure de probabilité. On calcule la probabilité pour cette séquence d'états d'être en adéquation avec le modèle :

$$d_t = \Pr(\mathbf{X}_{t-T+1}, \dots, \mathbf{X}_{t-2}, \mathbf{X}_{t-1}, \mathbf{X}_t) = q_{x_{t-T+1}} \prod_{\tau=t-T+2}^t p_{x_{\tau-1}x_\tau}$$

- La définition des instants de décision : les instants de décision sont les instants où on décide de déclencher le processus de détection de situations à risque. Ce processus est court-circuité dans le cas où le système reçoit des données alarmantes comme l'indication d'un arrêt cardiaque ou d'une chute brutale.

Indépendamment de ces situations extrêmes, les instants de décisions sont répartis au cours d'une journée en prenant en compte le rythme de vie de la personne. On peut faire varier cette répartition d'un jour à l'autre si la gravité de la situation de la personne augmente.

4.3.6 Décision sur la nature de la situation observée

La situation du patient à un instant donné est caractérisée par une mesure de probabilité représentant son adéquation au modèle d'évolution habituelle. Plus cette mesure est éloignée de la valeur maximale 1, plus la probabilité de gravité de la situation augmente. Cette mesure de probabilité est la base de la prise de décision sur le type de la situation observée.

Les principes de décision sont évalués à partir des premiers résultats d'expérimentation du modèle. L'étude des distances au modèle observées dans le cas d'évolutions habituelles et critiques permet d'apprécier les performances des différents principes de décision envisagés. Deux paramètres sont importants à spécifier dans le cas d'une prise de décision par seuillage sur la distance au modèle :

- Le nombre de classes correspondant à des situations de référence : ces situations correspondent aux types de situation envisagés, parmi lesquelles devra donc être classée chaque situation observée. Les situations de référence peuvent être ordonnées par gravité croissante. On peut par exemple choisir de considérer trois classes de situations : « situation habituelle », « situation préoccupante » et « situation critique ».
- Les seuils de décision : il s'agit de définir une méthode de prise de décision sur la situation à partir de la mesure de distance entre une situation observée et le modèle d'évolution : cet intervalle est $[0,1]$ puisqu'on a défini cette mesure par une mesure de probabilité. Chaque partition de cet intervalle correspond alors à une prise de décision sur l'état du patient : « habituel », « préoccupant » ou « critique » par exemple. Le nombre de seuils dépend du nombre de situations de référence considérées. Si on considère trois situations de référence, on doit alors déterminer deux seuils de décision. Ces seuils sont déterminés par l'expérimentation du système avec les données d'apprentissage, de validation et de test : on cherche à classer des situations dont on connaît la classe d'appartenance *a priori*.

4.3.7 Evaluation des performances du système

Pour tester les performances du système, il s'agit d'évaluer sa capacité à prendre de bonnes décisions. Il a été mis en évidence la nécessité d'identifier des indicateurs de performance pertinents (méthode COR par exemple). Ces tests de performance nécessitent de disposer d'une base de données de test. On a notamment besoin de séquences de vecteurs d'états du patient considérées comme habituelles, ainsi que de séquences représentant des scénarios d'évolution vers une situation critique.

4.4 Synthèse et schéma général du processus d'apprentissage

Le schéma de la Figure 9 résume les différentes phases du processus d'apprentissage dont l'objectif est la définition des paramètres du modèle d'évolution habituelle de la personne à domicile et du processus de décision. L'apprentissage est réalisé à partir d'ensembles de données de référence utilisés successivement :

- données d'apprentissage : successions d'états représentatives d'une évolution habituelle ;
- données de validation : successions d'états de même type que celles dédiées à l'apprentissage ;
- données de test : successions d'états évoluant vers une situation critique.

On retrouve dans le processus d'apprentissage les étapes du processus général de détection de situations critiques. Aux différentes étapes de ce processus s'ajoute une démarche de définition de certains paramètres du modèle :

intervalles de segmentation pour le codage des variables d'état ;
probabilités caractéristiques de la chaîne de Markov ;
seuils de décision pour le choix du type de situation rencontrée.

D'autres paramètres sont définis dès l'initialisation du processus :

Choix des variables d'état du patient ;
Définition des instants de décision ;
Choix de la taille de la fenêtre d'observation.

Le processus d'apprentissage, incluant le processus de détection, est joué avec tous les types de données d'apprentissage disponibles. Les données d'apprentissage sont à l'origine de la définition des paramètres, les autres sont destinées à la validation à l'exception de l'étape de détermination des seuils de décision pour laquelle tous les types de données sont nécessaires.

On peut ensuite comparer les performances du système en faisant varier les paramètres initiaux du modèle ainsi que les choix de paramètres effectués à chaque étape d'apprentissage. L'objectif est d'obtenir un système de décision dont les performances ont la meilleure adéquation possible avec les contraintes du projet : compromis entre le taux de fausses alarmes et le taux de non détections d'une situation critique.

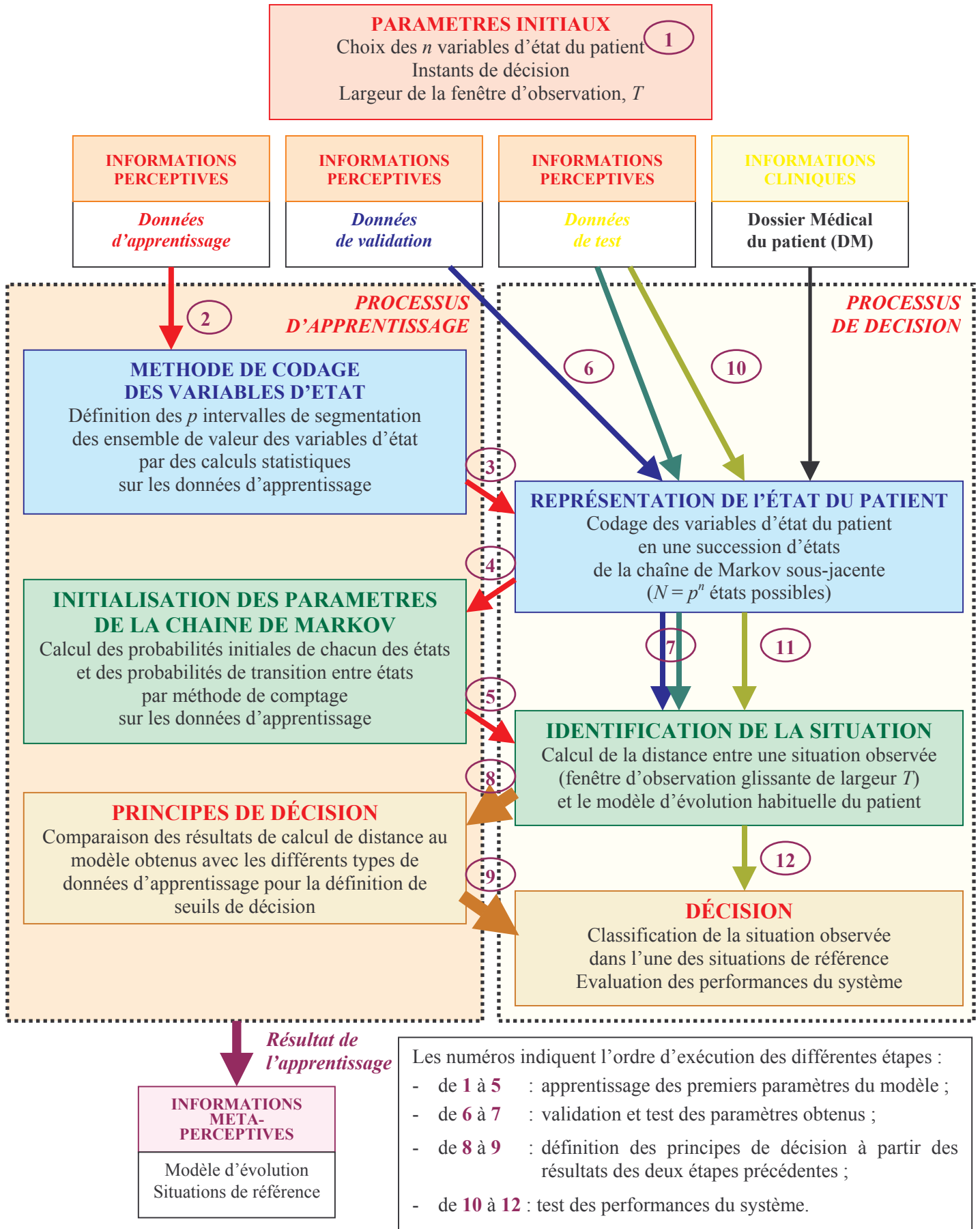


Figure 9. Les étapes du processus d'apprentissage du modèle

5 Application Java™ pour l'implémentation

Pour implémenter la construction et la mise en œuvre d'un modèle d'évolution habituelle du patient basé sur une chaîne de Markov, une application Java™ est développée. A partir de données représentatives du patient, elle doit permettre de définir la chaîne de Markov correspondante et de l'exploiter pour le calcul des distances d'une succession de vecteurs d'état au modèle d'évolution de référence.

5.1 Définition des fonctionnalités

L'application Java™ réalisée pour l'expérimentation répond à la fonctionnalité de création d'un patient et du modèle d'évolution associé et permet de calculer la distance d'une succession de vecteurs d'état par rapport au modèle ainsi défini.

Ses fonctionnalités graphiques sont décrites ci-dessous :

- Visualisation des particularités du patient (nom, prénom, date de naissance) et des caractéristiques de chacune des variables d'état (moyenne, écart-type, bornes inférieure et supérieure des intervalles de segmentation des valeurs);
- Visualisation des courbes de tendance de chacune des variables d'état considérées ;
- Visualisation des paramètres du modèle : vecteur des probabilités initiales des états et matrice des probabilités de transition entre états.

Par commodité vis à vis de leur exploitation, les distances calculées au modèle d'évolution lors des expérimentations successives ne sont pas affichées mais stockées dans des fichiers texte. Les courbes d'évolution sont ensuite tracées sous *Microsoft Excel*®.

Au niveau du cœur de l'application, les fonctionnalités principales mises en œuvre sont les suivantes :

- Récupération des données, c'est-à-dire d'une succession de vecteurs d'état du patient, à partir de fichiers texte.
- Définition des paramètres du codage des variables d'état, c'est-à-dire des intervalles de segmentation « habituel », « diminué » et « augmenté » :
 - calcul de la moyenne et de l'écart-type de l'échantillon d'entrée ;
 - définition des segments centrés sur la moyenne et de largeur égale à deux écart-types.
- Codage des vecteurs d'état afin que chaque variable d'état prenne ses valeurs dans l'ensemble {0, 1, 2}, à savoir {« habituel », « diminué », « augmenté »}.
- Définition des états de la chaîne de Markov à partir de la connaissance du nombre de variables d'état.
- Codage des vecteurs d'état en un numéro d'état de la chaîne de Markov.
- Initialisation des paramètres du modèle de Markov, c'est-à-dire des éléments suivants :
 - vecteur des probabilités initiales de chacun des états ;
 - matrice des probabilités de transition entre états.

Ces probabilités sont calculées par une méthode de comptage à partir d'une succession de vecteurs d'état.

- Calcul des probabilités d'adéquation au modèle (distance au modèle) à partir d'une succession de vecteurs d'état et selon la largeur de fenêtre d'observation définie initialement.

5.2 Processus d'exécution de l'application

L'exécution de l'application permet d'abord d'effectuer l'étape d'apprentissage du modèle, puis de réaliser différentes expérimentations selon les types des données fournies en entrée (succession de vecteurs d'état représentatifs d'évolutions habituelles ou critiques). Les éléments graphiques (profil du patient, évolution des variables d'état, caractéristiques du modèle d'évolution) sont ensuite affichés dans une fenêtre composée de plusieurs onglets, un par fonctionnalité graphique.

Les étapes successives de l'exécution de l'application sont présentées ci-dessous :

- Création du patient.
- Apprentissage du modèle :
 - Lecture des données d'apprentissage ;
 - Initialisation des paramètres du codage des variables d'état ;
 - Codage des données d'apprentissage en fonction de la délimitation des intervalles de segmentation ;
 - Codage de ces données en terme de numéro d'état dans la chaîne de Markov ;
 - Initialisation des paramètres de la chaîne de Markov ;
 - Calcul des probabilités d'adéquation au modèle à partir de ces mêmes données d'apprentissage
 - Ecriture des résultats dans un fichier texte.
- Validation du modèle :
 - Lecture des données de validation ;
 - Codage des données de validation en fonction de la délimitation des intervalles de segmentation ;
 - Codage de ces données en terme de numéro d'état dans la chaîne de Markov ;
 - Calcul des probabilités d'adéquation au modèle à partir de ces mêmes données de validation.
- Expérimentation du modèle :
 - Lecture des données d'expérimentation (évolution habituelle ou critique);
 - Codage des données d'expérimentation en fonction de la délimitation des intervalles de segmentation ;
 - Codage de ces données en terme de numéro d'état dans la chaîne de Markov ;
 - Calcul des probabilités d'adéquation au modèle à partir de ces mêmes données d'expérimentation.
- Affichage des éléments graphiques sous forme d'onglets dans une fenêtre graphique :
 - Onglet « Général » : profil du patient et caractéristiques de chacune des variables d'état. Le contenu de cette rubrique est affiché par défaut.
 - Un onglet par variable d'état portant le nom de la variable et permettant d'accéder aux courbes de tendance relatives au patient.
 - Un onglet « Modèle » contenant les paramètres du modèle d'évolution habituelle du patient : nombre d'états, vecteur des probabilités initiales, vecteur des probabilités de transition.

5.3 Modèle objet

Les fonctionnalités de l'application sont réalisées par un ensemble d'objets qui interagissent tout au long de la durée du processus d'exécution. Ces objets sont des instances de classes définies spécifiquement pour les besoins de l'application et constituant son modèle objet.

5.3.1 Description des classes

La globalité de l'application est une instance de la classe **Telecare** définissant la méthode principale *main* qui est la première invoquée. Elle contient une application, c'est-à-dire une instance *appli* de la classe **TelecareApplication**, et une interface graphique, instance de la classe **TelecareIUG**.

La classe **TelecareIUG** contient l'ensemble des méthodes nécessaires à l'affichage des éléments graphiques de l'application : les différents onglets et leur contenu. Elle inclut une référence de l'objet *appli*, de la classe **TelecareApplication**, et fait appel à certaines de ses méthodes pour récupérer les données à afficher.

La classe **TelecareApplication** contient un attribut *myPatient* de la classe **Patient**. Cette classe représente la personne télésurveillée à domicile. Elle gère l'ensemble des processus spécifiques au patient : l'apprentissage du modèle d'évolution et son expérimentation à partir d'un ensemble de données contenues dans des fichiers texte. L'objet *myPatient* de la classe **Patient** contient des attributs appartenant aux classes **StateVariable**, **StateVectors**, **StateVectorsManager** et **MarkovModel**.

La classe **StateVariable** gère les caractéristiques d'une variable d'état donnée.

La classe **StateVectors** permet de gérer un ensemble de données, c'est-à-dire une suite de vecteurs d'état du patient.

Les objets de la classe **StateVectorsManager** gèrent le calcul des probabilités d'adéquation des successions de vecteurs d'état au modèle d'évolution, en contrôlant notamment le déplacement de la fenêtre d'observation sur la suite des vecteurs d'état.

La classe **MarkovModel** représente le modèle de l'évolution habituelle du patient par une chaîne de Markov.

Enfin la classe **StateTranslation** permet de gérer le lien entre les caractéristiques des vecteurs d'état (nombre de variables d'état, intervalles de segmentation de leurs valeurs) et le modèle considéré, c'est-à-dire la manière dont on décide de définir les états de la chaîne de Markov à partir des vecteurs d'état. Elle réalise donc le lien entre les classes **StateVectors** et **MarkovModel**.

Une interface, **AppliConstants**, est utilisée pour la définition des constantes de l'application.

Les constantes de l'application ainsi que les attributs et les méthodes des différentes classes sont détaillés en Annexe H.

5.3.2 Interaction entre objets

Les paragraphes suivants décrivent les principales interactions entre objets lors des deux phases du déroulement de l'application : l'apprentissage du modèle et son expérimentation. On utilise pour cette description une représentation par certains diagrammes définis dans la notation **UML** (*Unified Modeling Language*) [45] développée par l'**OMG** (*Object management Group*) : les diagrammes de collaboration et les diagrammes de séquence.

Dans les diagrammes représentés, on ne considère que les interactions entre les objets créés spécifiquement pour l'application **Telecare**.

Des diagrammes plus détaillés spécifiant les arguments des messages échangés sont présentés en Annexe H.

5.3.2.1 Apprentissage du modèle : représentation par des diagrammes de collaboration

Les diagrammes de collaboration [45] montrent les interactions entre les objets en insistant particulièrement sur la structure spatiale statique qui permet la mise en collaboration d'un groupe d'objets. Les diagrammes de collaboration expriment à la fois le contexte d'un groupe d'objets (au travers des objets et des liens) et l'interaction entre ces objets (par la représentation des envois de messages). Une interaction est réalisée par un groupe d'objets qui collaborent en échangeant des messages.

Un diagramme de collaboration possède les caractéristiques suivantes :

- la syntaxe *:Classe* représente un objet anonyme du genre *Classe* ;
- les liens servent de support de transmission pour les messages ;
- les messages sont représentés le long des liens qui relient les objets, au moyen de flèches simples orientées vers le destinataire du message. Un message déclenche une action dans l'objet destinataire ;
- les flèches terminées par de petits cercles représentent les arguments des messages ;
- les différents messages sont numérotés pour indiquer l'ordre des envois ;
- une famille de liens, instances d'une même association, est représentée de manière condensée par un seul lien entre un objet et la famille d'objets qui lui est associée ;
- l'itération est indiquée par le caractère * placé devant le message ;
- les objets qui possèdent le flot de contrôle sont dits actifs. Un objet actif se représente par un rectangle dont la bordure est plus épaisse que celle des objets passifs.

Le diagramme de la Figure 10 est un diagramme simplifié représentant les principales interactions entre l'objet *myPatient* de la classe ***Patient*** et les objets auxquels il est lié lors du processus d'apprentissage du modèle d'évolution. La variable *n* représente le nombre de variables d'état considérées pour le patient. Huit messages successifs sont émis par l'objet du type *Patient*.

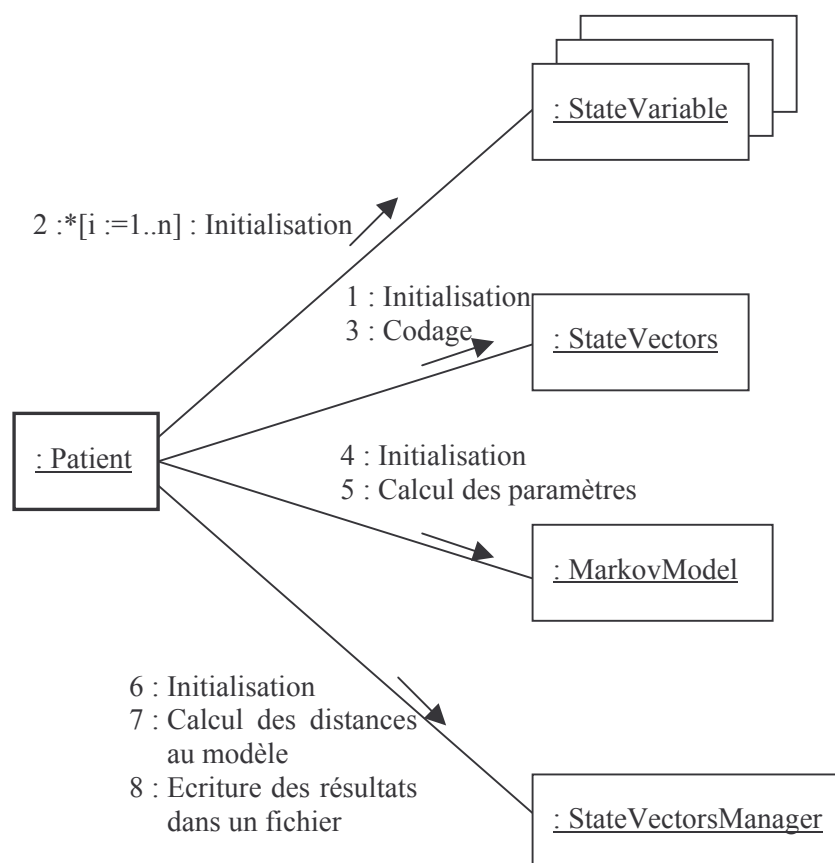


Figure 10. Représentation des interactions et de l'ordre des envois de message entre un objet de la classe **Patient** et les objets qu'il contient

Le diagramme de la Figure 11 représente les interactions détaillées entre tous les objets de l'application impliqués dans le processus d'apprentissage et d'expérimentation des paramètres du modèle.

Les variables qui apparaissent dans le diagramme sont définies ci-dessous :

- n : nombre de variables d'état considérées ;
- N : nombre d'états de la chaîne de Markov ;
- M : nombre de vecteurs d'état pris en compte pour l'apprentissage ;
- k : nombre maximal d'incrémentations possibles lorsqu'on fait glisser la fenêtre d'observation sur la succession de vecteurs d'état de taille M .

Les numéros indiquant l'ordre d'envoi des messages sont hiérarchisés à la manière des titres d'un document. L'objet :Patient est l'émetteur de chaque série de messages et chacun de ces messages porte donc un simple numéro x . Si l'objet récepteur doit lui aussi faire appel aux services d'un ou plusieurs autres objets, il numérote alors la série de messages émis par le numéro du message reçu suivi d'un numéro y , soit $x.y$. Et ainsi de suite à chaque niveau de collaboration.

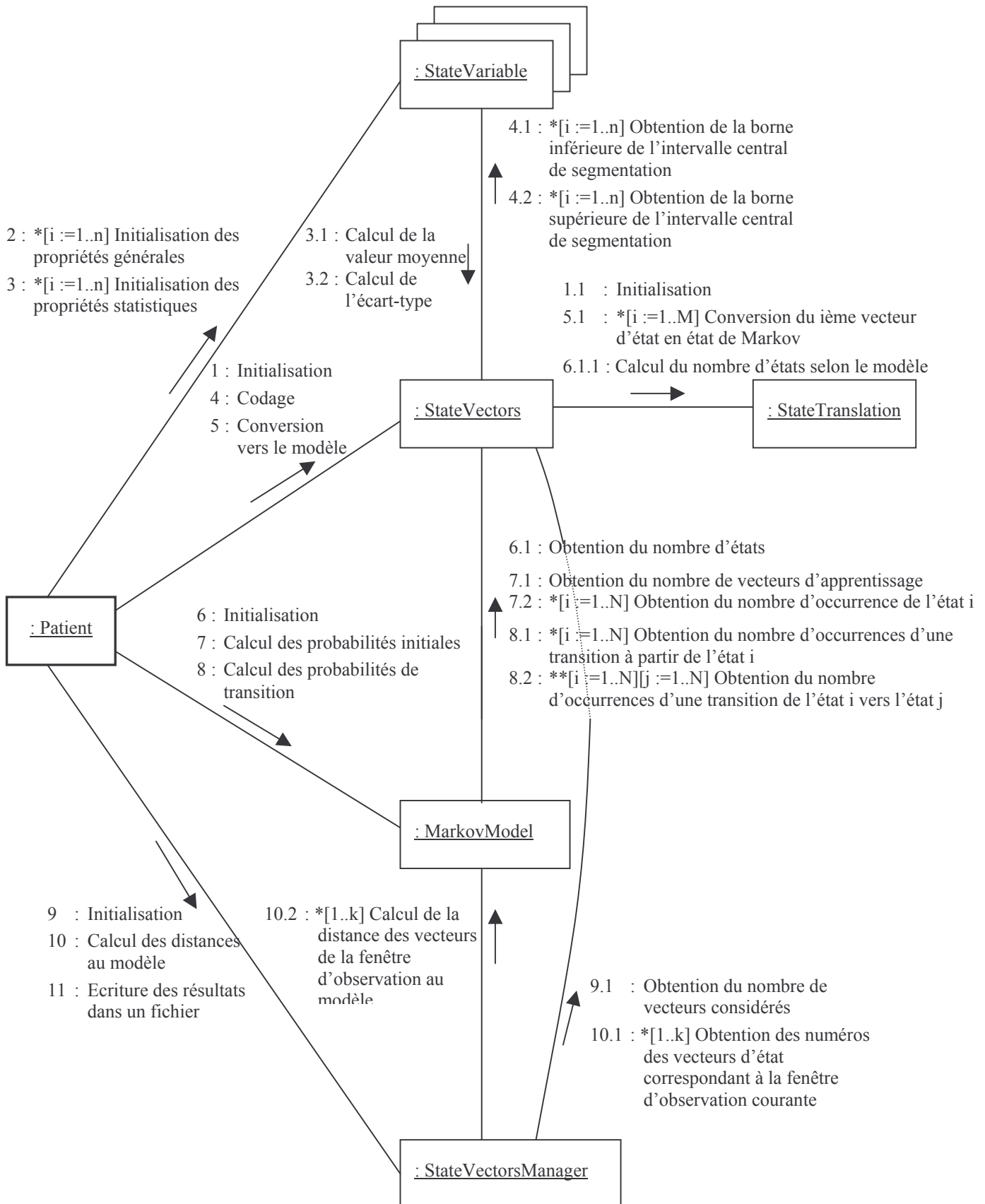


Figure 11. Représentation détaillée des interactions entre les objets de l'application

5.3.2.2 Expérimentation du modèle : représentation par un diagramme de séquence

Les diagrammes de séquence [45] montrent les interactions entre objets selon un point de vue temporel. Le contexte des objets n'est pas représenté de manière explicite comme dans les diagrammes de collaboration. La représentation se concentre sur l'expression des interactions.

Un diagramme de séquence possède les caractéristiques suivantes :

- la syntaxe : *Classe* représente un objet anonyme du genre *Classe* ;
- un objet est matérialisé par un rectangle et une barre verticale appelée ligne de vie des objets ;
- les objets communiquent en échangeant des messages représentés au moyen de flèches horizontales, orientées de l'émetteur du message vers le destinataire ;
- l'ordre d'envoi des messages est donné par la position sur l'axe vertical ;
- les périodes d'activité des objets, c'est-à-dire le temps pendant lequel un objet effectue une action, se représentent par des bandes rectangulaires placées sur les lignes de vies ;
- dans le cas des envois de message synchrones, le retour en fin d'exécution est implicite et ne nécessite pas de représentation particulière ;
- le cas particulier de l'envoi de messages récurifs se représente par un dédoublement de la bande rectangulaire.

Le diagramme de la Figure 12 représente les principales interactions entre l'ensemble des objets de l'application lors du processus d'expérimentation du modèle d'évolution du patient. On considère donc notamment que les paramètres du codage et de la chaîne de Markov sont déjà initialisés.

Dans ce diagramme, tous les envois de messages sont synchrones : l'émetteur d'un message est bloqué et attend que l'appelé ait fini de traiter le message. Les activités de plus bas niveau qui s'exercent au sein d'un objet ne sont pas représentées (cas d'un objet qui s'envoie lui-même un message), de même que la création d'objets temporaires (leur durée de vie est limitée au temps d'exécution d'une méthode d'un objet permanent de l'application).

Les formes des diagrammes de séquence reflètent indirectement les choix de structure. Le diagramme suivant présente un mode de contrôle centralisé à partir d'une instance de la classe *Patient*.

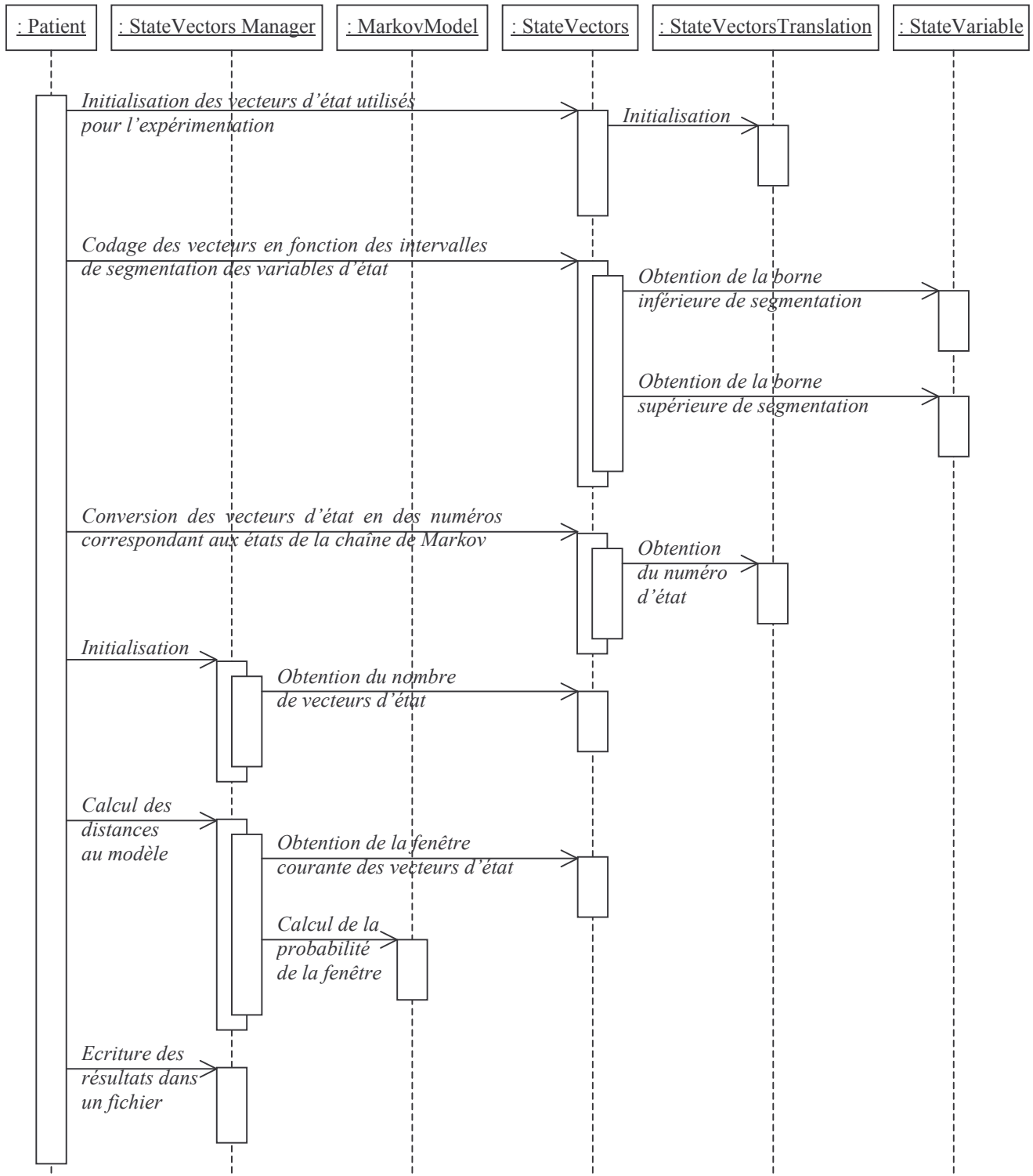


Figure 12. Représentation des principales interactions entre les objets de l'application lors du processus d'apprentissage des paramètres du modèle

6 Expérimentation du modèle

6.1 Introduction

Afin de déterminer si l'approche théorique présentée au paragraphe précédent est envisageable, on cherche à l'expérimenter sur un système de télésurveillance qui prend en compte un nombre restreint de variables d'état pour représenter la situation de la personne à domicile. Les expérimentations sont réalisées en utilisant une application Java développée spécifiquement. Cette application a été décrite au paragraphe précédent (cf. §5) et plus en détail en Annexe H.

L'objectif de cette démarche est de valider les choix théoriques et de soulever de nouveaux problèmes à prendre en compte dans le prochain incrément de la démarche de recherche. Il s'agit de construire le modèle et d'en évaluer ensuite les performances grâce à un ensemble de données de test (taux de non détections et de fausses alarmes). Cependant, même avec un nombre restreint de variables d'état, le manque de données pour la définition et le test du modèle oblige à la révision de ces objectifs (cf. §6.2.2). Il s'agit alors simplement de vérifier que les réponses de l'application aux différentes données fournies en entrée correspondent aux résultats attendus (cf. §6.3.1).

Dans la suite on commence par décrire les conditions de l'expérimentation : définition des paramètres, description des données expérimentales et des tests réalisés. Les résultats sont ensuite présentés et commentés pour aboutir à la définition d'un ensemble de contraintes à prendre en compte dans la prochaine étape de la recherche.

6.2 Conditions expérimentales

6.2.1 Paramètres initiaux

Il s'agit de définir pour chaque expérimentation réalisée l'ensemble des variables d'état prises en compte, les instants de décision ainsi que la largeur de la fenêtre d'observation de l'état du patient.

Contraint par les seules données dont on dispose, on choisit de considérer trois variables d'état pour représenter l'état du patient :

- la Pression artérielle Systolique (PS),
- la Pression artérielle Diastolique (PD),
- la Fréquence Cardiaque (FC).

On segmente les intervalles de variation de chacune de ces variables d'état en trois intervalles : « habituel », « diminué » et « augmenté ». L'intervalle central de variation (« habituel ») est défini à partir de l'estimation de la valeur moyenne m et de l'écart-type σ d'un échantillon représentatif d'une évolution habituelle du patient. On choisit un intervalle de largeur correspondant à deux écart-types et centré sur la moyenne : $[m-\sigma, m+\sigma]$.

La chaîne de Markov constituée de tous les états possibles du patient comporte ainsi $3^3 = 27$ états. Le vecteur des probabilités initiales de chacun des états est alors de dimension 27 et la matrice des probabilités de transition entre états de dimensions 27×27 , ce qui représente $27^2 = 729$ transitions possibles.

Les instants de décision sont ainsi les instants, non réguliers et s'étalant sur trois semaines environ, auxquels ont été mesurées les valeurs de ces trois variables dans l'ensemble des données disponibles. On n'associe donc aucune signification particulière à la répartition des instants de décision.

La largeur de la fenêtre d'observation est égale à 20 ou 30 instants de décision selon les expérimentations.

6.2.2 Données expérimentales

On dispose d'un ensemble de mesures cardiaques comprenant les données aux mêmes instants des pressions artérielles systolique et diastolique ainsi que de la fréquence cardiaque. Les données sont prises successivement à 92 instants. Les intervalles de temps entre chaque prise ne sont pas identiques. L'ensemble des valeurs est récolté sur une durée totale de 3 semaines environ.

On répartit les données dont on dispose en deux sous-ensembles de même cardinal égal à 46 valeurs destinés chacun à deux étapes distinctes d'apprentissage des paramètres du modèle :

- le premier sous-ensemble de données est utilisé pour l'apprentissage des paramètres du codage des variables d'état et des paramètres de la chaîne de Markov ;
- le second est utile à la validation du modèle.

On ne dispose *a priori* que de données représentatives d'une évolution habituelle du patient qu'il convient de compléter par des scénarios d'évolution vers des situations préoccupantes, voire critiques. On identifie trois types de tendance correspondant à une évolution inhabituelle des variables d'état du patient : l'augmentation, la diminution et l'augmentation des fluctuations des valeurs. Puisqu'on ne dispose d'aucune série de valeurs caractéristiques de ces évolutions, les données correspondantes sont calculées en ajoutant une tendance aux évolutions habituelles des valeurs des variables d'état.

On choisit de prendre en compte une variation d'amplitude δ sur une durée T . Les fonctions permettant le calcul des données de test sont présentées ci-dessous :

Soit $x(t)$ la fonction du temps t représentant la valeur d'une variable d'état X à l'instant t ;

Soit m la valeur moyenne de cette variable d'état ;

Soit $x_d(t)$ la fonction du temps t représentant la valeur de la variable d'état X à l'instant t suite à une tendance à la diminution des valeurs :

$$x_d(t) = x(t) - \frac{\delta}{T} \times t$$

Soit $x_a(t)$ la fonction du temps t représentant la valeur de la variable d'état X à l'instant t suite à une tendance à l'augmentation des valeurs :

$$x_a(t) = x(t) + \frac{\delta}{T} \times t$$

Soit $x_f(t)$ la fonction du temps t représentant la valeur de la variable d'état X à l'instant t suite à une tendance à l'augmentation des fluctuations des valeurs :

$$x_f(t) = x(t) + \text{signe}(x(t) - m) \times \frac{\delta}{T} \times t$$

Pratiquement, on choisit une variation d'amplitude $\delta = 10$ unités de mesure sur une durée $T = 46$ unités de temps correspondant à la durée des échantillons de test. Les unités de mesure sont les millimètres de mercure (mmHg) pour les pressions artérielles et le nombre de pulsations par minute (puls/min) pour la fréquence cardiaque.

Les tableaux de valeurs et les courbes correspondantes sont présentées en Annexe I.

On sait d'ores et déjà que l'on ne dispose largement pas de suffisamment de valeurs pour valider n'importe quelle hypothèse ou tirer des conclusions des résultats obtenus. A partir d'un échantillon de 46 valeurs, et donc de 45 transitions entre états du modèle, il est impossible d'estimer par une simple méthode de comptage les probabilités de transition entre les états alors qu'il en existe 729 possibles ! Les expérimentations permettent alors simplement de vérifier la pertinence des résultats obtenus, à savoir leur conformité avec ce qu'on pouvait en attendre. Nous décrivons dans les perspectives du projet quels peuvent être les axes de recherche à explorer pour faire face d'une part à la difficulté de récolter des données fiables, et d'autre part à l'explosion combinatoire du nombre d'états de la chaîne de Markov si un grand nombre de variables d'état sont prises en compte (cf. §7).

6.2.3 Processus d'expérimentation

L'expérimentation comporte deux phases principales pour la construction puis l'exploitation du modèle d'évolution d'une personne médicalement télésurveillée à domicile :

- l'utilisation de données d'apprentissage pour la construction du modèle : on suit alors le processus d'apprentissage présenté schématiquement au paragraphe 4.4 et représenté au niveau de l'application Java™ par un diagramme de séquence au paragraphe 5.3.2.2 ;
- l'exploitation du modèle ainsi défini à l'aide de données de test représentant différents types d'évolution possible pour le patient : on suit dans ce cas le processus général de détection présenté schématiquement au paragraphe 3.9 et représenté au niveau de l'application Java™ par un diagramme de collaboration 5.3.2.1.

Lors de la phase d'apprentissage, on expérimente trois types de modèles :

- le premier modèle (appelé modèle 1) est le modèle complet, c'est-à-dire que la chaîne de Markov sous-jacente est constituée de l'ensemble des états possibles du patient (27 états) ;
- dans le deuxième modèle (appelé modèle 2) on ne considère que 4 états possibles. Chaque état de ce modèle regroupe l'ensemble des états de la chaîne complète ayant le même nombre de variables d'état (entre 0 et 3) hors de l'intervalle défini comme « habituel » pour le patient ;
- dans le dernier modèle (appelé modèle 3) on regroupe au moment du codage des variables d'état les valeurs des intervalles « diminué » et « augmenté » dans une même classe « inhabituelle ». On n'a donc plus que 2 intervalles de segmentation et donc $2^3 = 8$ états possibles.

Rien ne valide cependant la pertinence de l'utilisation d'un modèle simplifié par rapport au modèle complet. Le manque de données ne peut en aucun cas justifier une simplification. Il faudrait pouvoir vérifier que les réponses du système de décision sont similaires quel que soit le type de modèle utilisé. On peut par exemple vérifier qu'on obtient les mêmes taux de non détections et de fausses alarmes en réponse à un ensemble de données de test. Dans ce cas, la simplification correspond uniquement à la suppression de bruit et non d'informations utiles, elle est donc justifiée.

Dans le cas de cette première expérimentation, on ne dispose pas de suffisamment de données pour une telle justification de simplification du modèle. On se contente donc d'observer que les résultats attendus correspondent dans tous les cas à ce qu'on pouvait en attendre.

Lors de la phase de test du modèle, on se réfère à trois types d'évolution inhabituelle pour une personne à domicile. Les deux premières évolutions considérées correspondent à une évolution inquiétante alors que la dernière tend à ramener le patient vers une situation « plus normale » en terme médical. Les caractéristiques des évolutions considérées sont décrites ci-dessous en terme de tendance de variation des valeurs de l'ensemble des variables d'état (informations fournies par L. Bajolle).

- Test 1 : Vieillessement

Une évolution caractéristique du vieillissement est généralement lente et peut s'observer significativement sur une période d'une année.

Pression artérielle Systolique (PS) : augmentation
 Pression artérielle Diastolique (PD) : augmentation
 Fréquence Cardiaque (FC) : augmentation des fluctuations

- Test 2 : Accident aigu

Cet accident correspond à un élargissement de la différentielle, c'est-à-dire de la différence entre les pressions artérielles systolique et diastolique. Cette évolution est très rapide (quelques heures).

Pression artérielle Systolique (PS) : augmentation
 Pression artérielle Diastolique (PD) : diminution
 Fréquence Cardiaque (FC) : stable

- Test 3 : Succès thérapeutique

Cette évolution peut être observée sur un sujet hypertendu suite à un traitement avec un médicament bêta-bloquant par exemple. Elle s'observe significativement à l'échelle d'un mois.

Pression artérielle Systolique (PS) : diminution
 Pression artérielle Diastolique (PD) : diminution
 Fréquence Cardiaque (FC) : diminution

Les données correspondant à ces différentes évolutions possibles sont calculées à partir des données brutes selon la méthode décrite au paragraphe 6.2.2.

6.3 Présentation des résultats

6.3.1 Caractéristiques statistiques des variables d'état

A partir de l'échantillon d'apprentissage, on calcule les propriétés statistiques (valeur moyenne et écart-type) de chaque variable d'état. Ces propriétés sont utiles à la délimitation des intervalles de segmentation des valeurs de ces variables destinés au codage (cf. 6.2.1) : $[m-\sigma, m+\sigma]$ est l'intervalle considéré comme « habituel », l'intervalle inférieur correspond alors à des valeurs « diminuées » et l'intervalle supérieur à des valeurs « augmentées ».

Les propriétés statistiques estimées à partir du tableau des 46 valeurs d'apprentissage disponibles (cf. Annexe I.1) sont présentées dans le Tableau 1. Les valeurs sont présentées arrondies à l'entier le plus proche. Cette approximation est effectuée après l'estimation de chacune des valeurs.

Tableau 1. Propriétés statistiques des variables d'état et limites de l'intervalle central de segmentation

	Valeur moyenne	Ecart-type	Borne inférieure	Borne supérieure
Pression artérielle systolique	127	10	117	137
Pression artérielle diastolique	78	9	70	87
Fréquence cardiaque	81	13	68	94

On détermine ainsi les paramètres du codage des variables d'état.

6.3.2 Paramètres de la chaîne de Markov

Les paramètres de la chaîne de Markov, c'est-à-dire les probabilités initiales de chacun des états et les probabilités de transition entre états, sont également estimés à partir des données d'apprentissage par une méthode de comptage.

Il a été souligné dans les paragraphes précédents l'impossibilité d'effectuer une bonne estimation à partir des données dont on dispose. Les résultats obtenus pour une chaîne de Markov à 27 états (modèle 1) sont néanmoins présentés dans les tableaux ci-dessous pour illustrer quelques caractéristiques que peuvent présenter les chaînes de Markov. Le vecteur des probabilités initiales est de dimension 27 ; la matrice des probabilités de transition est de dimension 27x27. Les numéros des lignes et des colonnes (non explicités) correspondent aux numéros des états de la chaîne de Markov (cf. 3.4.6 Principe de numérotation des états du modèle). Le signe « - » signifie que la probabilité estimée est nulle. Cette probabilité est remplacée lors du calcul de distance par la précision (P) choisie. Les autres probabilités sont alors arrondies à ce même niveau de précision. Dans le cas des données suivantes, on a choisi $P=10^{-3}$.

Vecteur des probabilités initiales des états (modèle 1)

0.391	0.13	0.065	0.065	-	0.022	0.065	0.022	0.022	0.022	-	0.022	0.043	-	0.022	-	-	0.022	0.022	-	-	0.022	-	0.043	-	-
-------	------	-------	-------	---	-------	-------	-------	-------	-------	---	-------	-------	---	-------	---	---	-------	-------	---	---	-------	---	-------	---	---

Matrice des probabilités de transition entre états (modèle 1)

0.5	0.111	0.056	-	-	-	0.056	-	-	-	-	-	0.056	-	0.056	-	-	0.056	0.056	-	-	0.056	-	-	-	-
0.167	0.167	-	0.333	-	-	0.167	-	-	-	-	0.167	-	-	-	-	-	-	-	-	-	-	-	-	-	-
0.667	-	-	-	-	-	-	-	-	0.333	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	1.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	0.333	-	0.333	-	-	-	-	0.333	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	1.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.0	-
-	1.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	1.0	-	-	-	-	-	-	-	-	-	-	-	-	-
0.5	-	-	-	-	-	0.5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	1.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	1.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
0.5	-	0.5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

- Un état (au niveau du vecteur des probabilités initiales) ou une transition (au niveau de la matrice des probabilités de transitions) ayant cette représentation correspond à une transition ou à une état initialement impossible.
- Une ligne de cette couleur correspond à un état « ligne » dont ne peut partir aucune transition. Le numéro de l'état dans la chaîne de Markov correspond au numéro de la ligne.
- Une colonne de cette couleur correspond à un état « colonne » qui ne peut jamais être atteint. Le numéro de l'état dans la chaîne de Markov correspond au numéro de la colonne.
- Un état « ligne » ainsi représenté est un état absorbant : il peut être atteint soit initialement, soit à partir d'un autre état, mais aucune transition ne permet d'en sortir.
- Un état ainsi représenté au niveau du vecteur des probabilités initiales est un état impossible : sa probabilité initiale est nulle et il ne peut jamais être atteint.

Figure 13. Paramètres de la chaîne de Markov représentant tous les états possibles du patient

L'illustration de quelques particularités que peut présenter une chaîne de Markov à partir des valeurs de paramètres présentées sur la Figure 13 ne permet en aucun cas de généraliser le fait que certains états sont absorbants ou impossibles par exemple. On pense d'ailleurs intuitivement qu'il ne doit pas y avoir d'états absorbants dans cette chaîne. On cherche à en effet à modéliser l'évolution habituelle d'un patient qui doit donc tendre à le ramener régulièrement dans l'état où la valeur de chaque variable d'état est habituelle (état 0).

Par contre, si on constate au terme d'une expérimentation valide que certains états sont impossibles, on peut alors simplifier la chaîne de Markov en supprimant ces états : les probabilités de transition vers cet état et la probabilité initiale de cet état est alors nulle et il n'est plus question de la remplacer par une valeur presque nulle (Précision P).

En utilisant un modèle simplifié (modèles 2 & 3), on obtient des résultats peut-être plus robustes car il y a beaucoup moins d'états dans la chaîne de Markov et ses paramètres sont mieux estimés à partir d'un faible ensemble de données. Rien ne justifie cependant ces simplifications. Il faudrait par exemple constater que les lignes et les colonnes de la matrice des transitions du modèle complet (modèle 1) sont identiques pour tous les états qui ont été regroupés pour constituer l'un des deux autres modèles proposés. Il est donc pour cela nécessaire de disposer d'une bonne estimation des paramètres du premier modèle.

La Figure 14 et la Figure 15 présentent respectivement les valeurs de paramètres obtenus pour la chaîne de Markov correspondant au modèle 2 (4 états) puis au modèle 3 (8 états).

Vecteur des probabilités initiales des états (modèle 2)

0.391	0.37	0.196	0.043
-------	------	-------	-------

Matrice des probabilités de transition entre états (modèle 2)

0.5	0.278	0.111	0.111
0.353	0.471	0.176	-
0.333	0.333	0.333	-
-	1.0	-	-

Figure 14. Paramètres de la chaîne de Markov à 4 états possibles pour le patient (modèle2)

Vecteur des probabilités initiales des états (modèle 3)

0.391	0.156	0.13	0.065	0.043	0.043	0.087	0.043
-------	-------	------	-------	-------	-------	-------	-------

Matrice des probabilités de transition entre états (modèle 3)

0.5	0.167	0.056	-	0.056	0.056	0.056	0.111
0.333	0.111	0.333	-	0.111	0.111	-	
0.5	0.167	0.167	0.167	-	-	-	
-	0.333	-	0.333	-	-	0.333	
-	0.5	-	0.5	-	-	-	
0.5	-	-	-	-	-	0.5	
0.5	0.25	0.25	-	-	-	-	
-	1.0	-	-	-	-	-	

Figure 15. Paramètres de la chaîne de Markov à 8 états possibles pour le patient (modèle3)

Dans les deux exemples précédents, le dernier état correspond à l'état « le plus inhabituel » pour la personne. Les valeurs de paramètres estimées correspondent alors bien aux résultats attendus pour la modélisation d'une évolution habituelle : plus les états considérés sont graves, plus les probabilités initiales et les probabilités de transition vers ces états sont faibles (valeurs des colonnes correspondant à ces états). Si l'un de ces états est atteint, l'état suivant sera très certainement un état « plus habituel » : les lignes de la matrice de transition correspondant à ces états « les plus inhabituels » ont des probabilités de transition importantes vers les autres états « plus normaux » et presque nulles vers les états les plus inhabituels.

6.3.3 Distance au modèle : résultats attendus

La mesure de distance d'une succession de vecteurs d'état du patient au modèle d'évolution habituelle correspond à un produit de probabilités : la probabilité d'occurrence du premier état et les probabilités de transition entre les états observés. Chacune de ces probabilités est très inférieure à 1 compte-tenu du nombre d'états considérés dans la chaîne de Markov. Leur produit est donc également très faible, d'autant plus quand on considère une fenêtre d'observation de plus en plus large.

En calculant les distances au modèle de la succession des vecteurs d'état utilisés pour l'apprentissage des paramètres du modèle, on s'attend à obtenir un comportement stationnaire de ces probabilités au cours du temps. Plus on considère une large fenêtre d'observation W , plus les distances calculées sont faibles puisqu'on augmente en même temps le nombre de termes du produit de probabilités correspondant (Figure 16).

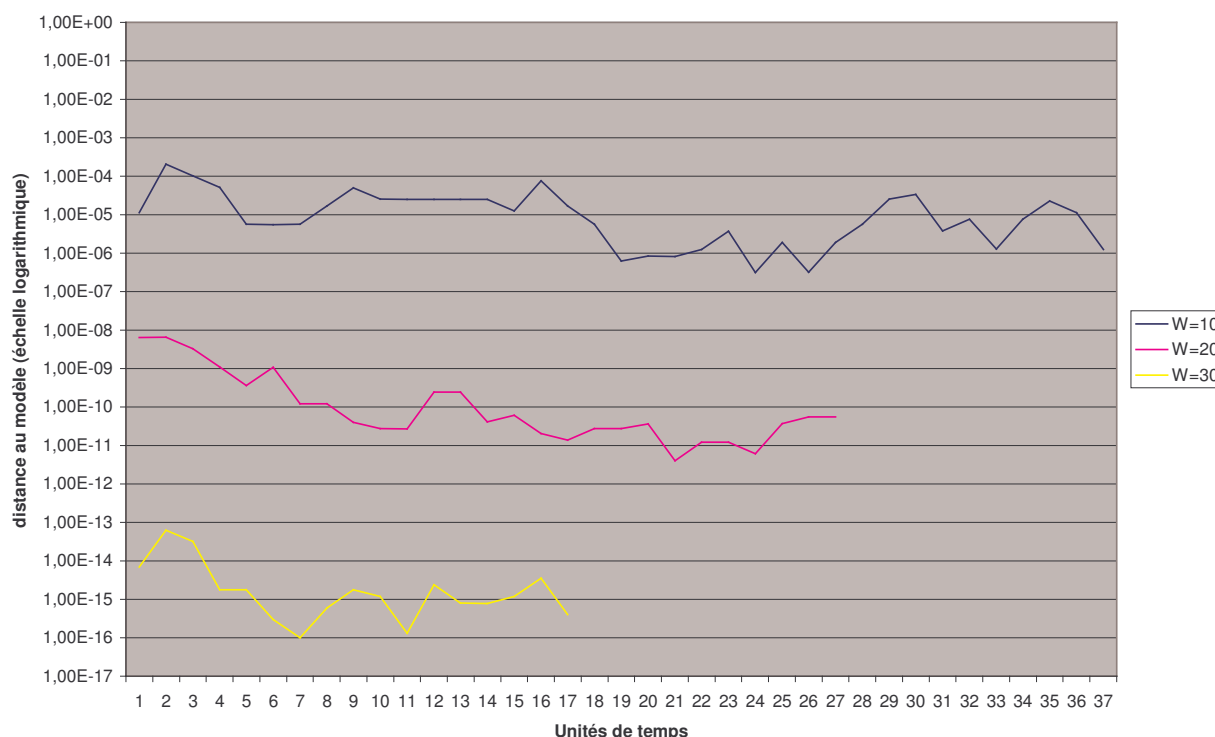


Figure 16. Evolution de la distance au modèle (27 états) de la succession des vecteurs d'état utilisés pour l'apprentissage en fonction de la largeur W de la fenêtre d'observation

Compte-tenu d'un apprentissage très limité du modèle, le système réagit ensuite de la même manière, qu'on lui fournisse des données représentatives d'une évolution habituelle (validation) ou non (tests d'évolution caractéristiques du vieillissement, d'un accident aigu ou d'un succès thérapeutique) : les probabilités d'adéquation au modèle sont beaucoup plus faibles que dans le cas d'une expérimentation du système avec les données d'apprentissage (Figure 17). Sur le graphique de la Figure 17, les expérimentations numérotées « 1 » correspondent à celles effectuées avec les données dérivées de l'apprentissage ; celles numérotées « 2 » sont effectuées à partir des données dérivées de la validation (cf. 6.2.2).

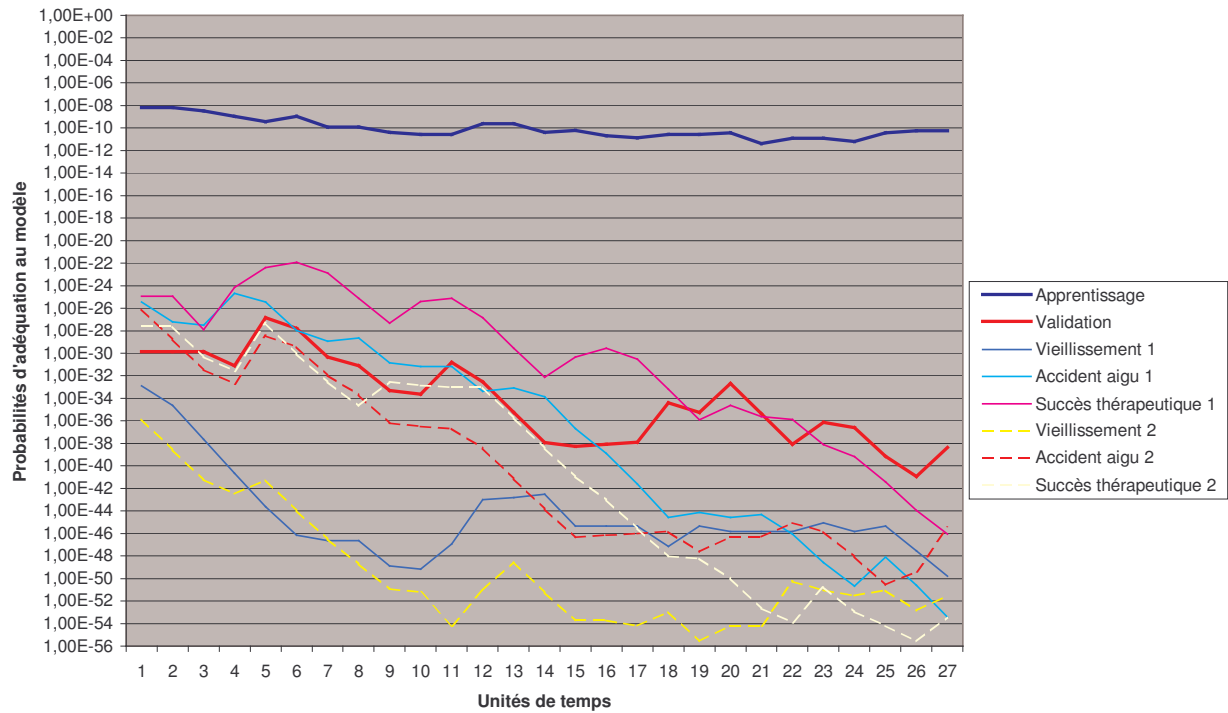


Figure 17. Evolution des probabilités d'adéquation au modèle 1 (chaîne de Markov à 27 états) des successions de vecteurs d'état correspondant aux données d'apprentissage, de validation et de test (Taille de la fenêtre $W=20$ – Précision $P=10^{-3}$)

On obtient le même type de résultats en expérimentant le modèle 3 à 8 états (Figure 18).

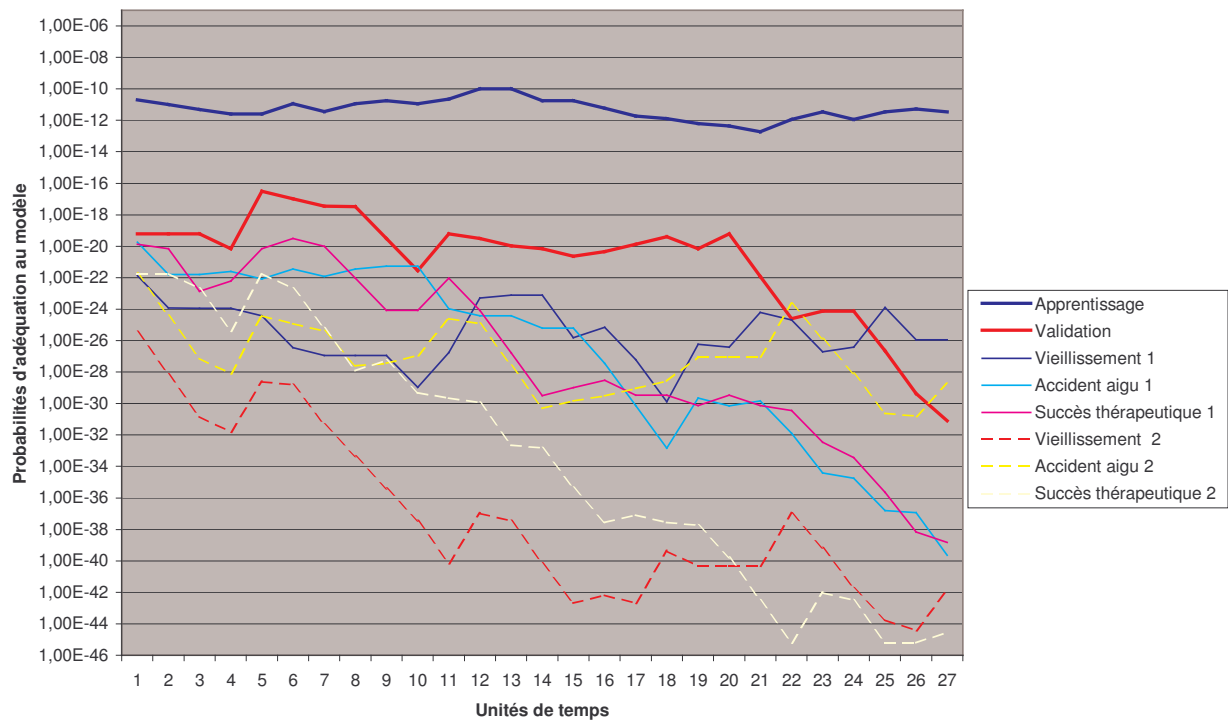


Figure 18. Evolution des probabilités d'adéquation au modèle 3 (chaîne de Markov à 8 états) des successions de vecteurs d'état correspondant aux données d'apprentissage, de validation et de test (Taille de la fenêtre $W=20$ – Précision $P=10^{-3}$)

Avec la mise en œuvre du modèle à 4 états (modèle 2), on note tout de même une différence par rapport aux résultats précédents pour ce qui concerne les données de validation : les distances au modèle calculées sont stationnaires et proches des résultats obtenus avec les données d'apprentissage, même si elles restent tout de même plus faibles. Les expérimentations avec les différentes données de test semblent tendre vers une diminution des probabilités d'adéquation au modèle. Les résultats obtenus sont présentés sur le graphe de la Figure 19.

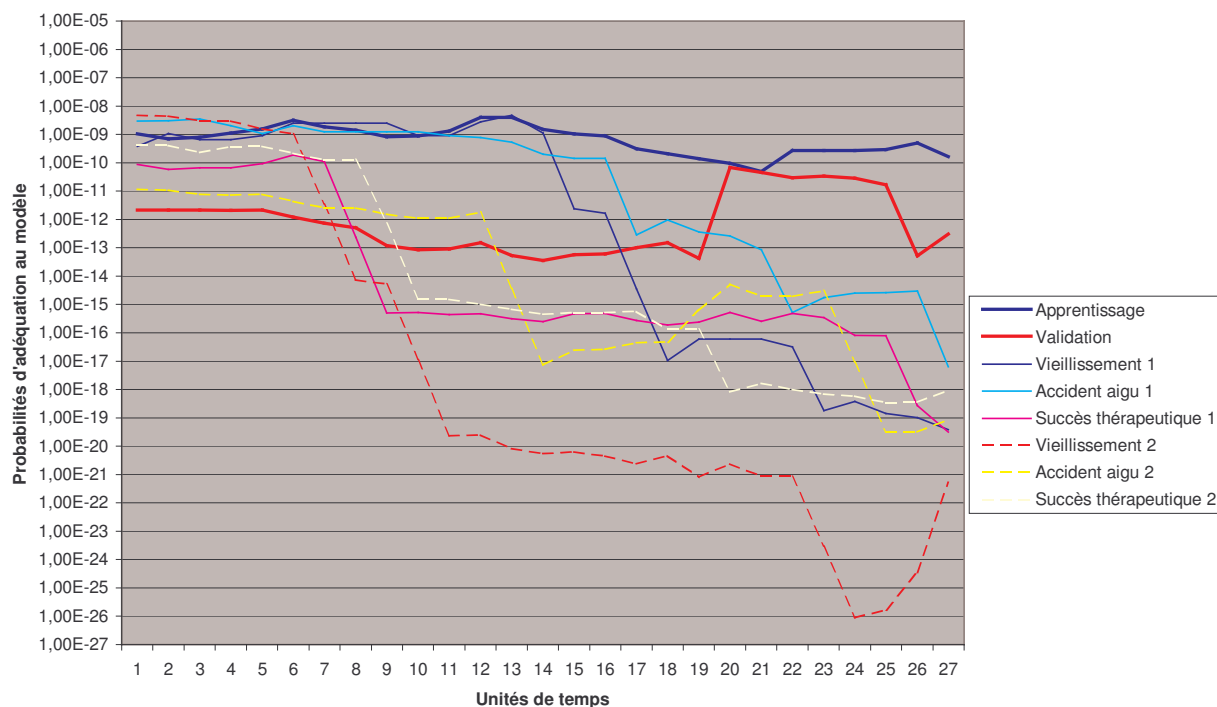


Figure 19. Evolution des probabilités d'adéquation au modèle 2 (chaîne de Markov à 4 états) des successions de vecteurs d'état correspondant aux données d'apprentissage, de validation et de test (Taille de la fenêtre $W=20$ – Précision $P=10^{-3}$)

Il ne sert cependant à rien d'étudier plus en avant ce modèle tant que la simplification proposée n'est pas justifiée. Cela autorise tout de même à penser qu'on pourrait obtenir de meilleurs résultats avec une chaîne à 27 états en disposant de plus de données d'apprentissage.

En choisissant pour la même expérimentation (modèle 2) une précision de 10^{-5} au lieu de 10^{-3} , la précision correspondant à la valeur de substitution des probabilités nulles des paramètres de la chaîne de Markov, on constate que les résultats obtenus sont similaires avec les données d'apprentissage et de validation, mais la tendance à l'éloignement vis à vis du modèle pour ce qui concerne les données de test est plus marquée (courbes en pointillés pour $P=10^{-5}$ par rapport aux courbes en traits pleins pour $P=10^{-3}$ sur le graphe de la Figure 20). Cet écart augmenté permet alors de mieux détecter les éloignements au modèle puisqu'il ne change pas la distance calculée avec les données caractéristiques d'une évolution habituelle du patient.

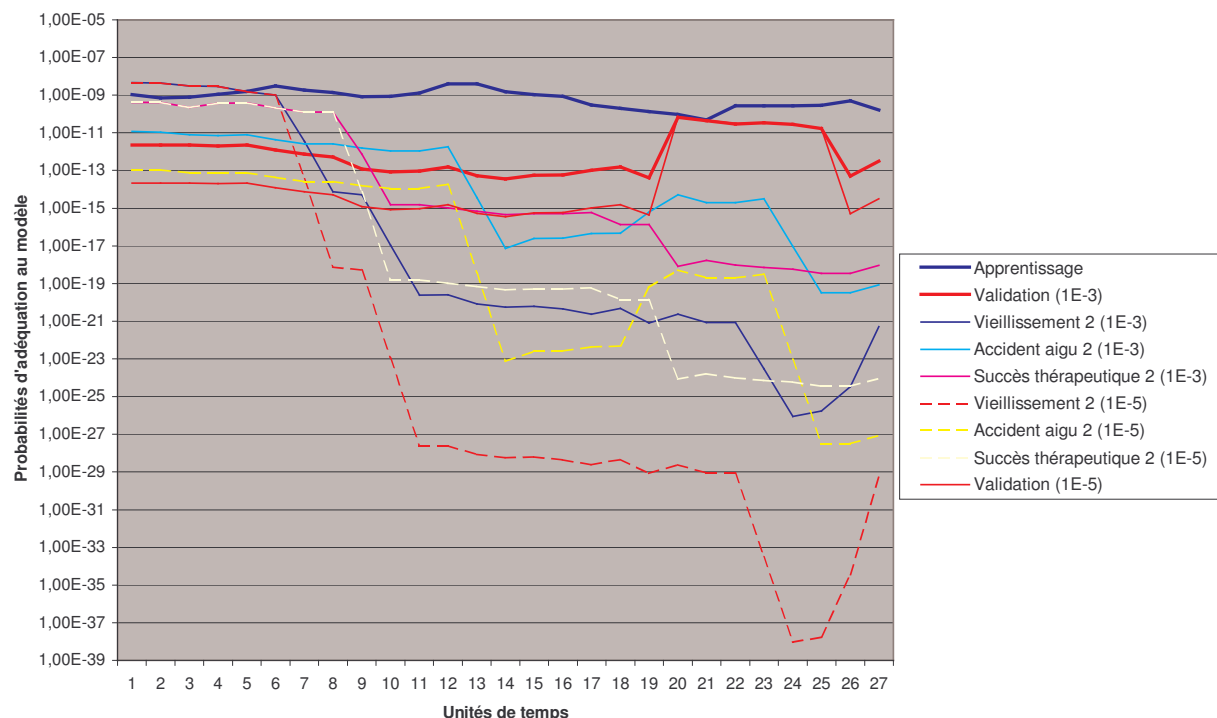


Figure 20. Evolution des probabilités d'adéquation au modèle 2 (chaîne de Markov à 4 états) des successions de vecteurs d'état correspondant aux données d'apprentissage, de validation et de test avec une précision $P=10^{-3}$ et $P=10^{-5}$ (Taille de la fenêtre $W=20$)

6.4 Commentaires

L'ensemble des graphiques présentés ci-dessus ne peut malheureusement pas être exploité compte-tenu des problèmes liés à l'apprentissage des paramètres du modèle. On ne dispose pas de suffisamment de données pour paramétrer le modèle complet et de ce fait on ne peut pas non plus justifier une quelconque simplification de ce modèle. On ne peut que constater l'adéquation des résultats avec les tendances attendues.

C'est cependant à partir de ce genre d'expérimentation que doivent être effectués les tests des principes de décision quant à la situation de la personne à domicile : critique ou non par exemple. Si on décide de placer un seuil sur les probabilités d'adéquation au modèle, cela revient à tracer sur le graphe un ensemble de lignes horizontales au niveau de chacun de ces seuils. On teste ensuite un ensemble d'évolutions habituelles et critiques. Si pour une évolution habituelle la courbe des distances au modèle passe sous le seuil, on observe une fausse alarme. Si une courbe correspondant à une évolution critique est au-dessus du seuil, il s'agit d'une non-détection. On voit apparaître ici la notion du temps de détection : à partir de quel intervalle de temps peut-on considérer qu'une situation critique qui reste au-dessus du seuil correspond à une non détection ?

Le manque de données pour l'expérimentation pose tout de même sérieusement la question des moyens à mettre en œuvre pour en obtenir suffisamment par la suite. En ne considérant que trois variables d'état, on obtient déjà une chaîne de Markov complète à 27 états. Que penser alors de la taille qu'atteindra cette chaîne, et de là l'ensemble d'apprentissage nécessaire, si on est simplement amené à doubler le nombre de variables d'état prises en compte ?

Idéalement, pour récolter des données fiables, il faudrait pouvoir observer une personne évoluant à son domicile pendant une longue période et relever régulièrement les valeurs des variables d'état considérées.

Il faudrait cependant des milliers de valeurs pour un seul patient afin d'espérer construire un modèle fiable. On aimerait ensuite tester le principe de modélisation sur plusieurs patients. Comment réussir à obtenir un si large ensemble de données ?

Ces remarques laissent à penser qu'il faudrait réussir à simplifier le modèle complet sans pour autant avoir besoin de l'expérimenter pour justifier les simplifications. Des axes de recherche pour tenter d'y parvenir sont proposés dans le paragraphe suivant consacré aux perspectives de ces premiers travaux.

7 Perspectives

Les paragraphes suivants présentent les perspectives de la recherche compte-tenu des réflexions engagées suite aux premières recherches et expérimentations d'un modèle de détection. Il s'agit notamment d'envisager les moyens de collecter des données fiables pour l'expérimentation, ainsi que les méthodes à exploiter pour tenter de valider certaines simplifications à apporter au modèle. La collecte de données fiables semble être un point sensible sur ce projet. Si on peut considérer un modèle composé d'un nombre réduit d'états par rapport au modèle complet, on aura d'autant moins besoin de données d'apprentissage, de validation et de test pour paramétrer le système de détection de situations critiques.

7.1 Collecte de données relatives au patient

Pour la collecte et l'exploitation de données médicales, deux types de méthodes sont couramment utilisés : les études rétrospectives et les études prospectives.

Une étude rétrospective consiste à récupérer des données dans les dossiers médicaux des patients suite à plusieurs années de suivi médical. Cette pratique est cependant peu fiable compte-tenu du fait que l'ensemble des données ainsi rassemblées est très souvent incomplet : l'absence de la notification d'un symptôme par exemple est interprétée dans l'étude comme « symptôme inexistant » alors que celui-ci peut simplement ne pas avoir été considéré comme digne d'intérêt à un moment donné par l'auteur du rapport. Cela a été le cas par exemple pour ce qui concerne les études de ce type menées pour tenter de lier l'hypertension et l'apparition de céphalées. De simples maux de tête d'un patient ne sont pas forcément notifiés dans un rapport tant qu'aucun autre symptôme n'est constaté. Si par contre on note une hypertension chez ce patient, on aura tendance à être plus vigilant vis à vis de tout autre signe qui peut paraître intéressant dans ces nouvelles circonstances. Une étude rétrospective peut alors amener à conclure qu'une hypertension apparaît en même temps qu'une céphalée, ce qui n'est pas exact.

Si on cherche à appliquer une méthode rétrospective dans le cadre du projet SIC-HIS pour collecter des données, on est de toutes façons confronté au problème de la connaissance des paramètres environnementaux liés aux différentes valeurs : température ambiante, activité de la personne à domicile, etc. On ne dispose dans ce cas que de données « hors contexte », ce qui ne nous permet pas d'exploiter toute la richesse de l'environnement de l'habitat intelligent permettant de prendre en compte des informations de nature variée. Le seul ensemble de données identifié comme pouvant être pertinent est celui rassemblant les informations liées aux blocs opératoires : des informations environnementales telles que la température ambiante, la ventilation sont enregistrées en même temps que des mesures cliniques sur le patient. Cependant, on se place alors dans le contexte d'une situation inhabituelle pour le patient.

Les études prospectives consistent, à partir de la définition d'un ensemble de critères à considérer pour une étude donnée, à observer l'évolution de ces différents critères sur un échantillon et à mener ensuite l'étude à partir de l'ensemble des données collectées. Cela revient à observer plusieurs personnes caractéristiques de la population à laquelle s'adresse le projet de télésurveillance dans différents appartements. Cette méthode est beaucoup moins biaisée que la précédente, sauf dans le cas où les échantillons sont mal identifiés (ne considérer par exemple que des personnes de moins de 30 ans dans le cadre de la télésurveillance médicale à domicile). Cette méthode est cependant difficile et longue à mettre en œuvre.

Au niveau du projet HIS, on se place *a priori* plutôt dans le cadre d'une étude prospective. L'idéal serait qu'un habitat spécialement équipé, par exemple celui de la faculté de médecine de Grenoble, soit réellement occupé par une personne qui accepterait de se soumettre aux contraintes de la télésurveillance : prise régulière du poids, de la tension, etc. Il faudrait même pouvoir observer plusieurs personnes ayant des profils différents parmi la population à laquelle s'adresse ce projet, c'est-à-dire les personnes âgées particulièrement. Il serait donc adapté d'entreprendre la démarche d'installation de l'équipement nécessaire à la télésurveillance médicale dans une maison de retraite par exemple. Cette démarche est cependant longue et coûteuse. On met ainsi en évidence les difficultés liées à la collecte de données fiables et pertinentes.

7.2 Généralisation du codage des variables d'état

Tel qu'il a été décrit jusqu'à présent, le principe proposé pour la modélisation de l'évolution d'une personne à domicile et la détection de situations critiques est très spécifique au patient télésurveillé. Cela signifie qu'il faut reconsidérer le modèle et donc rejouer le processus d'apprentissage des paramètres à chaque nouveau patient. Cette démarche est lourde surtout compte-tenu de la grande quantité de données nécessaires pour cette tâche.

Une idée est donc de pouvoir généraliser les paramètres estimés à l'ensemble des patients. Cela signifie qu'il n'est plus nécessaire de procéder à une nouvelle estimation du modèle chaque fois qu'une personne choisit la télésurveillance à domicile, mais également qu'on peut considérer un ensemble plus large de données autorisées lors de cette estimation : on peut étendre la collecte à une population de référence, c'est-à-dire à un échantillon de plusieurs patients représentatifs de la population concernée par la télésurveillance à domicile (personnes âgées notamment).

La généralisation du modèle d'évolution habituelle d'une personne à domicile à un ensemble de patients est valide si toute la spécificité du patient peut être codée dans le choix des variables d'état et la délimitation des intervalles de segmentation des valeurs de ces variables. Il s'agit de placer la variance inter-sujets dans une direction orthogonale à la direction du codage. Le modèle de la chaîne de Markov peut alors être considéré comme indépendant du patient et représente les évolutions médicales possibles d'une certaine normalité.

La possibilité d'envisager cette généralisation du modèle passe par une recherche bibliographique importante, en particulier sur les méthodes de codage déjà exploitées des différents paramètres médicaux considérés.

7.3 Simplification du modèle

Un autre axe de recherche à exploiter est celui des méthodes de simplification de la chaîne de Markov sous-jacente au problème de décision. Il s'agit de rechercher s'il est possible de justifier certaines réductions du nombre d'états à considérer dans la chaîne de Markov par rapport au modèle complet dans lequel on considère tous les états possibles du patient suite au codage des variables d'état. Deux points d'entrée sont possibles : soit on réduit le nombre de composantes d'un vecteur d'état, soit on réduit directement le nombre d'états par des regroupements ou des suppressions de certains d'entre eux.

7.3.1 Réduction du nombre de composantes des vecteurs d'état du patient

La réduction du nombre de composantes se justifie par une analyse de données sur l'ensemble des variables d'état. Il s'agit notamment d'étudier les corrélations entre ces variables, les qualités de discrimination de chacune d'entre elles vis à vis du problème de détection de situations critiques. On détermine ainsi les meilleurs paramètres ou variables d'état à considérer dans la modélisation en définissant de nouveaux paramètres synthétiques et en ne retenant que les variables les plus discriminantes. Si aucune simplification de ce type n'est identifiée, c'est que chaque variable d'état considérée porte des informations utiles à la détection des situations critiques qui ne se retrouvent dans aucune autre variable. Il faut alors toutes les prendre en compte.

L'observation de la fréquence cardiaque à partir des données disponibles laisse par exemple supposer l'existence de certaines corrélations entre cette variable médicale et peut-être d'autres variables environnementales. La répartition des valeurs semble en effet être bi-modale : aucune valeur n'est située sur la moyenne et l'ensemble des valeurs est réparti autour de deux modes. Cette évolution est présentée sous forme d'une courbe (Figure 21) et d'un histogramme (Figure 22).

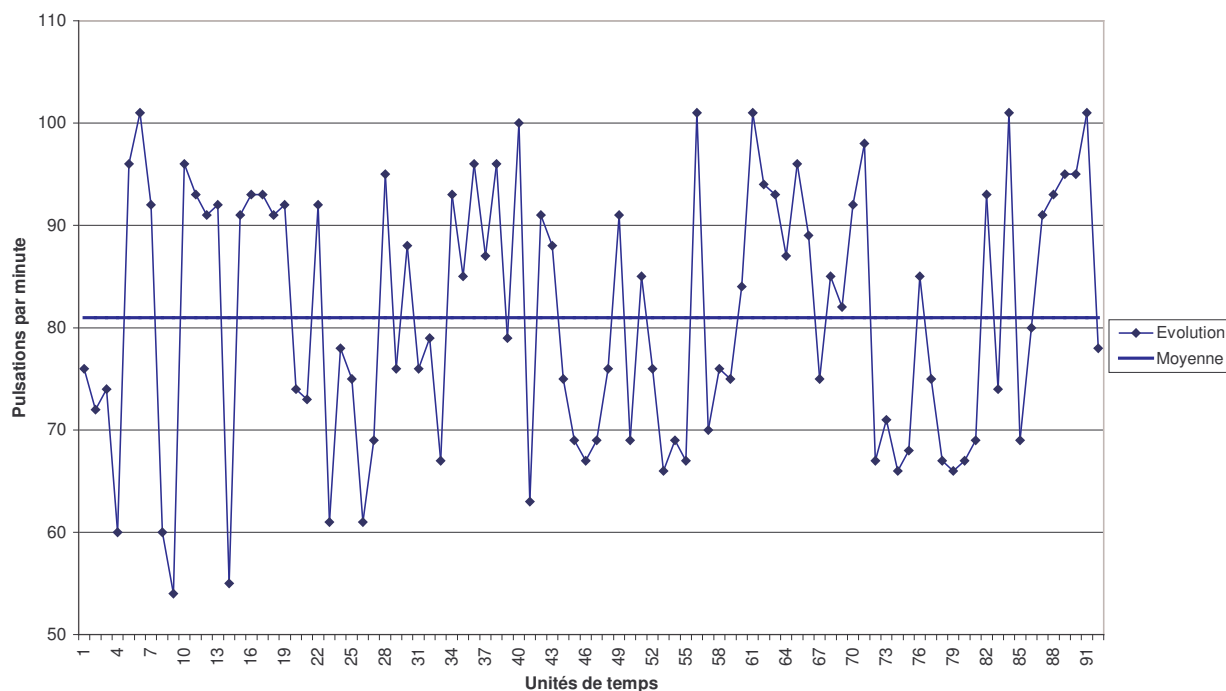


Figure 21. Evolution habituelle de la fréquence cardiaque d'un patient

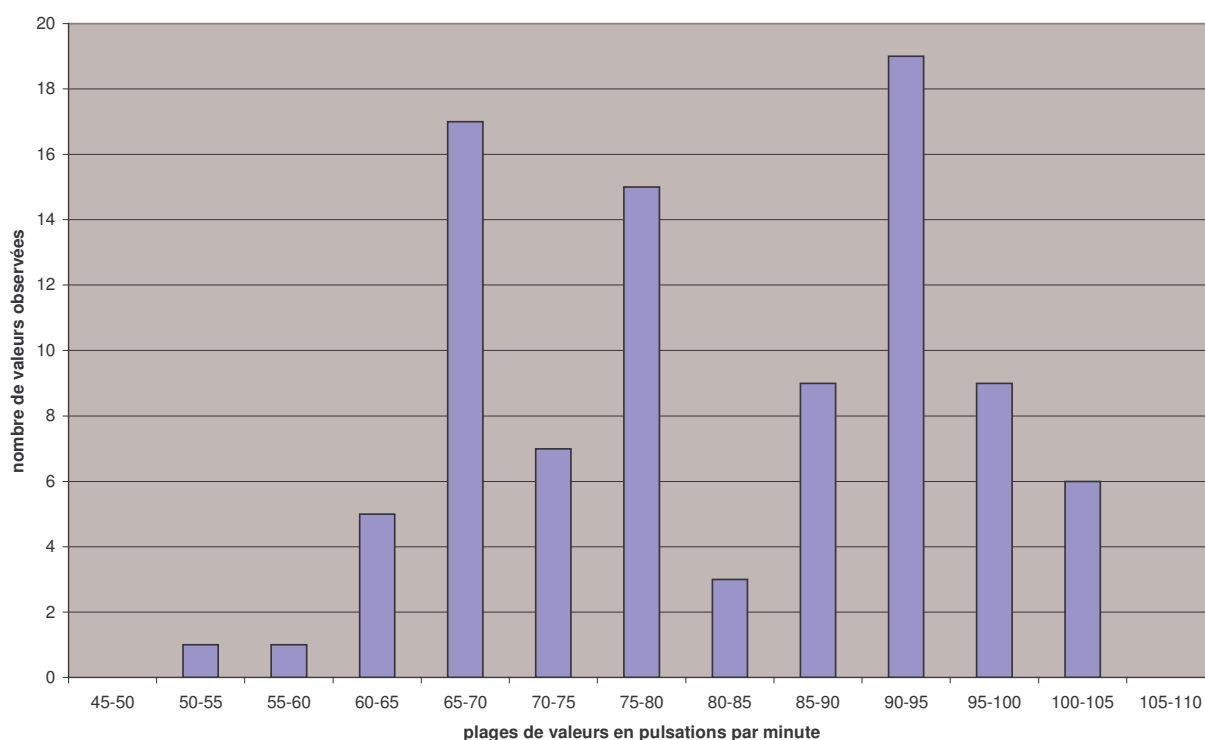


Figure 22. Répartition des valeurs de fréquence cardiaque pour une même patient

La fréquence cardiaque est certainement très corrélée à la température ambiante et à l'activité de la personne avant la prise de tension. S'il est justifiable de normaliser les valeurs de fréquence cardiaque par ces deux paramètres environnementaux par exemple, on enlève ainsi du bruit de la simple donnée de cette fréquence. Si par ailleurs les paramètres environnementaux ainsi considérés n'influencent sur aucune autre variable d'état, il est possible de les supprimer du vecteur d'état puisque toute l'information qu'ils portent est intégrée dans le modèle suite à la normalisation de la fréquence cardiaque.

7.3.2 Suppression d'états de la chaîne de Markov

Une fois la chaîne de Markov définie, c'est-à-dire une fois identifiées les variables d'état les plus pertinentes vis à vis du problème de détection, l'étude des paramètres estimés de cette chaîne peut justifier de nouvelles simplifications. Il s'agit notamment de repérer les états impossibles que l'on peut alors supprimer de la chaîne de Markov. Une autre simplification envisageable consiste à fusionner les états sans incidence sur la décision.

7.4 Apprentissage des paramètres du modèle

Une recherche bibliographique sur les méthodes d'amélioration de l'estimation des paramètres du modèle est intéressante. La technique de Monte Carlo par exemple est couramment utilisée dans les systèmes à temps discret pour faire face au problème d'explosion combinatoire du nombre d'états des modèles. Il s'agit d'étudier dans quelle mesure cette technique pourrait s'appliquer au modèle considéré dans le cas de la télésurveillance médicale à domicile.

7.5 Intervention des paramètres environnementaux

Les vecteurs d'état du patient sont *a priori* composés d'un ensemble de variables d'état médicales ou environnementales. On peut penser que ces données peuvent apporter beaucoup de bruit dans le modèle dans la mesure où elles ne sont corrélées que dans certaines circonstances aux observations effectuées sur la santé d'une personne : dans le cas d'une grande activité du patient ou d'une chaleur excessive dans l'appartement par exemple. On peut alors se poser la question du meilleur moment à considérer pour la fusion de ces données environnementales. On a déjà souligné que ces données pouvaient être prises en compte au niveau du codage des variables d'état (normalisation de la fréquence cardiaque par exemple). Il pourrait ensuite être plus efficace de sortir ces paramètres du modèle et de s'en servir plutôt pour définir au besoin des pondérations sur les distances calculées au modèle d'évolution. Il est ainsi possible de diminuer la sensibilité du système.

7.6 D'autres points de vue...

Au niveau du codage des variables d'état, le principe de délimitation stricte d'intervalles permettant de qualifier les valeurs semble un peu drastique. Les données médicales ne sont pas attachées à une sémantique précise et on aimerait pouvoir les qualifier en utilisant la notion de sous-ensembles flous. Cette idée attrayante remet néanmoins sûrement en questions le processus de décision mis en œuvre. Il faut en effet s'intéresser alors aux automates d'états flous et aux principes de décision à partir de ce type de modèle.

Une autre idée à garder en tête est celle du modèle de Markov caché dont la chaîne d'états modélise l'évolution d'un patient parmi l'ensemble des situations de référence : « situation habituelle », « situation préoccupante », « situation critique » et « décès ». Ce modèle est très plaisant car il semble bien en adéquation avec la problématique posée et est facilement compréhensible. La difficulté réside cependant dans la collecte des données nécessaires à l'apprentissage de ce modèle puisqu'on ne sait même pas définir quel peut être l'ensemble des évolutions critiques possibles.

8 Conclusion

Ce stage m'a donné l'occasion d'une première véritable expérience dans le milieu de la recherche. Cette étape était très importante pour moi puisqu'elle s'est déroulée dans la continuité d'une réorientation professionnelle suite à deux années passées dans l'industrie dès l'obtention du diplôme d'ingénieur. Peu satisfaite par ce contexte de travail, j'ai alors décidé de reprendre les études en DEA afin de me diriger vers les domaines de l'enseignement et de la recherche. Cette démarche m'a confortée dans l'idée que ce milieu correspond beaucoup mieux à mes aspirations. Je peux ainsi poursuivre les études de 3^{ème} cycle en toute confiance vis à vis des mes perspectives de carrière.

Dès les premières semaines de travail j'ai pris conscience de l'ampleur que peuvent très vite prendre les recherches bibliographiques. Aucun axe de recherche particulier n'ayant été défini au départ dans mon sujet, je me suis trouvée face à un très large ensemble de domaines théoriques à explorer. Le système étudié s'est avéré devoir se diviser en plusieurs étapes présentant chacune des adéquations à différents domaines de recherche. De plus, chaque article étudié référence lui-même un ou plusieurs autres articles qui peuvent sembler particulièrement intéressants. On assiste ainsi rapidement à une explosion combinatoire du nombre de références peut-être pertinentes vis à vis du sujet ! Afin de répondre à mon souci d'expérimenter un modèle avant la fin du stage, il fallait pourtant que je réussisse à cibler un axe de recherche particulier. Cette étape de choix d'une orientation consiste à restreindre au fur et à mesure l'étendue des domaines explorés en fonction des objectifs et des contraintes du projet. J'ai apprécié lors de cette démarche d'avoir l'opportunité de rencontrer des personnes extérieures au projet mais intéressées par des domaines communs de recherche. Les choix effectués lors de cette première immersion dans le projet ne sont cependant pas exhaustifs. D'autres domaines pouvant également présenter un intérêt ont été cités dans le document et pourront être explorés ultérieurement.

Une fois identifié un modèle pertinent de décision sur la situation d'une personne à domicile, j'ai été confrontée au problème stratégique de collecte de données fiables pour l'expérimentation. Il s'agit également d'en disposer en quantité suffisante pour permettre la validation du modèle et l'évaluation des performances du système de décision. En début de stage, l'ensemble des capteurs situés dans l'habitat équipé de la faculté de médecine de Grenoble n'était pas en état de fonctionnement et ne permettait alors pas cette collecte de données. C'est cependant le seul moyen dont on dispose actuellement pour observer une personne dans le contexte de la télésurveillance médicale à domicile. La démarche longue et coûteuse d'installation de l'équipement nécessaire à la télésurveillance médicale dans une maison de retraite par exemple n'était bien sûr pas non plus envisageable dans le cadre de ce stage. Il ne restait alors plus qu'à placer l'étude dans un cadre rétrospectif et à chercher un moyen de récupérer des données fiables et pertinentes. Finalement Lionel BAJOLLE, qui effectue sa thèse de médecine au sein du laboratoire, m'a proposé un ensemble de mesures cardiaques dont il disposait pour que je puisse expérimenter l'application développée pour construire et exploiter le modèle identifié. Cet ensemble de valeurs largement insuffisant ne m'a cependant pas permis d'aboutir à des résultats sur la validité du modèle et les performances du système de décision.

La poursuite de ce travail de recherche est donc fortement confrontée à ce problème de collecte de données fiables et pertinentes vis à vis du sujet. On a également mis en évidence dans le paragraphe consacré aux perspectives la nécessité d'une démarche de généralisation du codage des variables d'état et de simplification du modèle en fonction des particularités du contexte. Il s'agit notamment d'effectuer une analyse fine des données prises en compte dans le processus de décision. L'étape suivante de recherche nécessite donc un large ensemble de données dont on ne dispose pas actuellement. Il faut alors envisager tous les moyens possibles pour disposer rapidement de données fiables et pertinentes. La mise en place d'un véritable environnement de télésurveillance médicale pour la collecte de données n'étant pas réalisable à court terme, d'autres démarches doivent être envisagées. Il s'agit par exemple de se renseigner auprès d'hôpitaux, de maisons de retraites ou d'autres centres de recherche sur les possibilités d'obtention de ces ensembles de valeurs.

Bibliographie

1. Cauville, M., *Diagnostic, soins et prévention par la télémédecine : explications de J. Demongeot*, in *Sciences et technologies*, 1999, p. 32-34.
2. *Centre de Ressources sur le Télétravail et les Téléactivités en Auvergne (CRTTA) : Sommaire de la Télémédecine*, 2000, Association pour le Développement des Moyens Informatiques de la Région Auvergne (ADMIRA) ; <http://www.admira.asso.fr/crtta/home.htm>.
3. Faure, H., Rossignol, G., *La médecine à distance*, 1996, Direction des Hôpitaux du ministère de l'emploi et de la solidarité, Paris ; <http://www.sante.gouv.fr/htm/pointsur/telemed/80.htm>.
4. *Déclaration de l'Association Médicale Mondiale sur la surveillance des soins à domicile, la télémédecine et l'éthique médicale*, 1992, Association Médicale Mondiale (AMM) ; http://www.wma.net/f/policy/17-35_f.html.
5. Frawley, W. J., Piatetsky-Shapiro, G. and Matheus, C. J., *Knowledge discovery in databases : An overview*. AI Magazine, 1992, **13**(3): p. 57-70.
6. Lavrac, N. *Data Mining in Medicine: Selected Techniques and Applications*. in *Second International Conference on The Practical Applications of Knowledge Discovery and Data Mining*. 1998. London.
7. Fayyad, U. and Uthurusamy, R., *Data mining and knowledge discovery in databases (editorial)*. Communications of the ACM, 1996, **39**(11): p. 24-26.
8. Fayyad, U., Piatetsky-Shapiro, G. and Smyth, P., *The KDD Process for Extracting Useful Knowledge from Volumes of Data*. Communications of the ACM, 1996, **39**(11): p. 27-34.
9. French, S., *Decision theory*, 1986, Ellis Horwood.
10. Michie, D., Spiegelhalter, D. and Taylor, C., *Machine learning, neural and statistical classification*, 1994, New York, Ellis Horwood.
11. Pearl, J., *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, 1988, San Mateo, California, Morgan Kaufmann.
12. Degoulet, P. and Fieschi, M., *Traitement de l'information médicale : Méthodes et applications hospitalières*, Manuels Informatiques Masson - Entreprise, 1991, Masson. p. 320.
13. Zelic, I., Kononenko, I., Lavrac, N. and Vuga, V., *Induction of decision trees and Bayesian classification applied to diagnosis of sport injuries*. Journal of Medical Systems, 1997, **21**(6): p. 429-444.
14. Cao, C.G. and Leong, T.Y., *A learning approach to knowledge acquisition for supporting Markov therapy decision making*, in *Working notes of the AAAI Spring Symposium on Artificial Intelligence in Medicine: Applications of Current Technologies*, 1996, p. 11-15.
15. Saporta, G., *Probabilités, Analyse des Données et Statistique*, 1990, Paris, Technip. p. 493.
16. Kononenko, I., *Inductive and Bayesian learning in medical diagnosis*. Applied Artificial Intelligence, 1993, **7**: p. 317-337.
17. Steimann, F., *On the use and usefulness of fuzzy sets in medical AI*. Artificial Intelligence in Medicine, 2001, **21**: p. 131-137.
18. Steimann, F. and Adlassnig, K.-P., *Fuzzy Medical Diagnosis*, in *EH Ruspini, PP Bonissone, W Pedrycz (eds) Handbook of Fuzzy Computation*. 1998, IOP Publishing: Bristol.
19. Lavrac, N., Keravnou, E. and Zupan, B., *Intelligent Data Analysis in Medicine and Pharmacology*, 1997, Kluwer.
20. Clark, P. and Niblett, T., *The CN2 induction algorithm*. Machine Learning, 1989, **3**(4): p. 261-283.

21. Quinlan, J.R., *Induction of decision trees*. Machine Learning, 1986, **1**(1): p. 81-106.
22. Lavrac, N. and Dzeroski, S., *Inductive Logic Programming: Techniques and Applications*, 1994, New York, Ellis Horwood.
23. Dasarathy, B.V., *Nearest Neighbor (NN) Norms: NN Pattern Classification Techniques*, 1991, Los Alamitos, CA, IEEE Computer Society Press.
24. Aha, D., Kibler, D. and Albert, M., *Instance-based Learning Algorithms*. Machine Learning, 1991, **6**: p. 37-66.
25. Leong, T.Y., *A new methodology for clinical decision analysis over time: Theory and practice*, in *Working Notes of the AAAI Spring Symposium on Artificial Intelligence in Medicine: Applications of Current Technologies*, 1996, p. 89-93.
26. Leong, T.Y., *Dynamic decision modeling in medicine: A critique of existing formalisms*, in *Proceedings of the 17th Annual Symposium on Computer Applications in Medical Care*, 1993, p. 478-484.
27. Cestnik, G., Kononenko, I. and Bratko, I., *A Knowledge-Elicitation Tool for Sophisticated Users*. Progress in Machine Learning, 1987: p. 31-45.
28. Bratko, I. and Kononenko, I., *Learning diagnostic rules from incomplete and noisy data*, in *AI Methods in Statistics (Phelps, B., ed.)*. 1987, Gower Technical Press: London.
29. Cao, C., Leong, T.Y., Leong, A. P. K. and Seow, F.C., *Dynamic decision analysis in medicine: A data driven approach*. International Journal of Medical Informatics, 1998, **51**(1): p. 13-28.
30. Dean, T., *Decision-Theoretic Planning and Markov Decision Process*, A tutorial presented at the Summer Institute on Probability and Artificial Intelligence, 1994, Corvallis, Oregon. p. 31.
31. Magni, P., Quaglini, S., Marchetti, M., Barosi, G., *Deciding when to intervene: a Markov decision process approach*. International Journal of Medical Informatics, 2000, **60**(3): p. 237-253.
32. Sutton, R. S., Barto, A. G., *Reinforcement Learning: An Introduction*, Adaptive Computation and Machine Learning, 1998, Cambridge, MA, MIT Press. p. 432.
33. Littman, M. L., Dean, T. L., Kaelbling, L. P., *On the Complexity of Solving Markov Decision Problems*, in *Eleventh International Conference on Uncertainty in Artificial Intelligence*, 1995, Montréal, Québec, Canada, p. 394-402.
34. Puterman, M. L. and Feinberg, E. A., *Markov decision processes : Discrete stochastic dynamic programming*. SIAM Review, 1996, **38**(4): p. 689-?
35. Sundaresh, S., Leong, T.Y. and Haddawy, P., *Supporting multi-level multi-perspective dynamic decision making in medicine*, in *Proceedings of the 1999 AMIA Annual Fall Symposium*, 1999, AMIA, p. 161-165.
36. Lane, T., *Hidden Markov Models for Human/Computer Interface Modeling*, in *Proceedings of the IJCAI-99 Workshop on Learning About Users*, 1999, p. 35-44.
37. Lundin, E., Jonsson, E. *Some Practical and Fundamental Problems with Anomaly Detection*. in *proceedings of the Fourth Nordic Workshop on Secure IT systems (NORDSEC'99)*. 1999. Kista, Sweden.
38. Ghosh, K. A., Schwartzbard, A., Schatz, M., *Learning Program Behavior Profiles for Intrusion Detection*, in *Proceedings of the 1st USENIX Workshop on Intrusion Detection and Network Monitoring*, 1999, Santa Clara, California, p. 51-62.
39. Lee, W. and Stolfo, S., *Data mining approaches for intrusion detection*, in *Proceedings of the 7th USENIX Security Symposium*, 1998, San Antonio, Texas, p. 26-29.
40. Kemeny, J. G. and Snell, J. L., *Finite markov chains*, The university series in undergraduate mathematics, 1960, Van Nostrand. p. 210.

41. McEliece, R.J., *The theory of information and coding*, 1977, Addison-Wesley Publishing Company.
42. Aycard, O., *Architecture de contrôle pour robot mobile en environnement intérieur structuré*, in *Département de formation doctorale en informatique*, 1998, Université Henri Poincaré - Nancy 1, Nancy, France, 230 pages.
43. Ye, N. *A Markov chain model of temporal behavior for anomaly detection*. in *Proceedings of the 2000 IEEE, Workshop on Information Assurance and Security*. 2000. United States Military Academy, West Point, NY.
44. Metz, C.E., *Basic Principles of ROC Analysis*, in *Seminars in Nuclear Medicine*, 1978, vol. **8**, p. 283-298.
45. Muller, P.A., *Modélisation objet avec UML*, Eyrolles ed, 1997, Paris. p. 421.
46. Nemo, A., *La Télé médecine : Faire voyager les informations plutôt que le malade*, in *Le Journal du Téléphone*, 1994, p. 4.
47. *Information sur l'hypertension artérielle*, 2000, Hypertension On-line ; <http://www.hypertension-online.com/>.
48. *La plongée avec l'amiral : le sang et la circulation*, 2000 ; <http://plongee.amiral.free.fr/formation/niveau4/circuln4.htm>.
49. *Biotechnologies et Sciences Médico-Sociales : Apport des ressources numériques pour l'étude de l'appareil cardio-vasculaire*, 2000, EDUCNET ; <http://www.educnet.education.fr/bio/appcardi.htm>.
50. *Mesure de la pression artérielle*, 2000, BMédia ; <http://www.snv.jussieu.fr/bmedia/ATP/pressio.htm>.
51. Delpuech, F., *Un indicateur simple et objectif*, in *La Recherche - Dossier : Obésité, l'épidémie*, 2000, vol. **330**.
52. Swets, J.A., Pickett, R.M., *Evaluation of Diagnostic Systems : Methods from Signal Detection Theory*, Series in Cognition and Perception, 1982, New-York, Academic Press.
53. Davies, R., *Introductory Java for scientists and engineers*, 1999, Addison-Wesley. p. 294.

Annexe A Télémédecine

A.1 Définition

La télémédecine représente l'utilisation des **Nouvelles Technologies de l'Information et de la Communication (NTIC)** dans le secteur médical [1]. Elle médiatise l'acte médical en interposant un outil de communication entre les médecins ou entre un médecin et son patient. Ainsi, elle remet en cause une partie de la pratique médicale, mais représente un enjeu considérable. Elle va permettre l'amélioration des conditions de vie et la prise en charge de milliers de personnes.

A l'origine, dans les années soixante, la télémédecine a été développée pour répondre au problème d'isolement géographique de certaines populations de l'Amérique rurale principalement (Texas, Géorgie), mais aussi du nord de la Norvège. Au début des années quatre-vingt-dix, de nouvelles recherches ont été menées en Amérique pour tenter de trouver, par ce type d'organisation, une solution à la difficulté d'accès aux centres de soins spécialisés [2].

La télémédecine a aujourd'hui trouvé de nombreux champs d'applications, et se décline en différents termes [2, 3]:

- **Télé-assistance** : Aide thérapeutique directe apportée au malade.
- **Télé-consultation** : Consultation médicale à distance.
- **Téléchirurgie/
Télémanipulation** : Manipulation de matériel médical (instruments chirurgicaux) contrôlée à distance par le praticien sur le patient.
- **Télédiagnostic** : Echange de données entre équipes médicales dans une démarche d'aide au diagnostic.
- **Télé-expertise** : Avis donné à distance, en temps réel ou différé, par un expert ou un médecin à partir de la transmission de données médicales multimédia.
- **Télé-imagerie** : Envoi et consultation d'images médicales à distance.
- **Télé-opération** : Opération conduite à distance, notamment avec l'usage d'outils chirurgicaux.
- **Téléenseignement/
Téléformation** : Utilisation de l'outil informatique en particulier pour l'aide à la formation médicale continue des médecins : contacts professionnels via le réseau, consultation des informations médicales (banques de données, imagerie, suivi d'études épidémiologiques et d'essais clinique multicentriques, cours de formation).
- **Télésurveillance** : Surveillance médicale à distance de patients qui peuvent être à leur domicile.
- **Réseau de télémédecine** : Généralisation d'un réseau à l'ensemble d'un système de santé national ou supra-national pour la gestion de l'information médicale entre professionnels de santé : transmission des dossiers...

A.2 Les apports de la télémédecine

Le développement de la télémédecine est gouverné par des raisons pragmatiques de réduction des coûts, d'accroissement de l'efficacité des soins, ou de partage des connaissances. Pour chaque acteur en télémédecine, les avantages de ce type d'organisation sont nombreux [46].

Pour les praticiens, il s'agit de développer une plus grande coopération entre les différents réseaux du milieu médical : ville-hôpital, généraliste-spécialiste, public-privé. L'idée est de créer des passerelles de communication, d'information et de transmission de savoir. La profession médicale se trouve donc

réorganisée par le développement actuel de la télémédecine et le sera encore progressivement par ses évolutions :

- Le médecin pourra par exemple accéder facilement à l'avis d'un expert lors d'une problématique qui dépasse son domaine de compétences.
- Plusieurs médecins pourront aussi confronter leurs diagnostics en temps réel sur une pathologie particulière.
- Un autre avantage au niveau de l'amélioration de la transmission des connaissances pourrait être la possibilité qu'aurait un médecin d'accéder en temps réel à une banque de données faisant état de l'historique des interventions des acteurs en télémédecine, accompagné de conseils théoriques et d'avis empirique sur un problème donné.

L'intérêt des pouvoirs publics pour la télémédecine est directement lié à la réduction des coûts de santé publique. Ce type d'organisation permet en effet de mieux contrôler les dépenses de santé. Les gains de productivité concernent les économies de transport, mais aussi la mise en œuvre des soins. En effet, la télémédecine limite les déplacements des patients et du personnel médical, diminue la redondance des soins grâce à l'accès distant au dossier médical du patient à tout moment, etc.

Pour les patients, la télémédecine permet de répondre au problème d'isolement géographique en assurant l'égalité d'accès aux soins quel que soit le lieu d'habitation des personnes. Les petits centres hospitaliers souffrent en effet du manque d'équipement, mais surtout d'une pénurie de médecins compétents.

Par ailleurs, si on considère le cas particulier de la surveillance à distance, la télémédecine répond au besoin d'autonomie, de sécurité et de dignité de patients souhaitant rester à leur domicile. Elle s'inscrit alors dans la dynamique des alternatives à l'hospitalisation et devrait encore favoriser le développement des soins à domicile.

A terme, la télémédecine pourrait également aussi agir en faveur du transfert mondial de connaissances médicales, et améliorer l'aide aux pays en voie de développement ou émergents.

Le développement de la télémédecine intéresse également beaucoup certains secteurs médicaux pour lesquels elle serait parfois l'unique solution d'intervention pour l'apport de soins. Il s'agit par exemple de la médecine maritime, de la médecine sportive, de l'armée qui voient là un moyen d'aider à distance les marins, sportifs en zone isolée, soldats, spationautes, etc.

A.3 Les freins

Le développement de la télémédecine se heurte à des problèmes d'ordre psychologique, juridique, ou éthique et à des réticences de la part des différents acteurs.

De manière générale, certains restent encore sceptiques quant à l'intérêt intrinsèque de la télémédecine. Médecins et patients craignent notamment qu'elle porte atteinte à la liberté d'exercice, au secret médical, et conduise finalement à une déshumanisation de la relation entre le médecin et son patient.

De nombreuses réticences sont dues à la nécessité de changement de la structure organisationnelle du monde hospitalier et médical (changement des habitudes de travail, peur de l'informatique, manque de temps,...). En France, la lourdeur administrative du secteur de la Santé fait de la télémédecine un domaine « contraint » qui ne peut connaître que des évolutions lentes.

Ces nouvelles pratiques médicales soulèvent de nombreux problèmes éthiques et juridiques. L'authenticité et la confidentialité des données du patient doivent être garanties à tout moment. L'utilisation de l'outil informatique pour la consultation, le transfert et la sauvegarde des informations concernant les patients ne doit pas nuire à leur confidentialité et à leur fiabilité. Par ailleurs dans le cadre de la télé-expertise par exemple, l'avis consultatif d'un confrère, même s'il s'agit d'un acte de prudence, ne dégage rien le médecin traitant de sa propre responsabilité vis-à-vis de son patient. Cela le rend alors plus vulnérable en terme de responsabilité civile.

Une autre crainte est celle de la fuite des compétences médicales des centres de soins les plus isolés. La délocalisation d'opérations médicales est en effet accompagnée du risque de regroupement des meilleurs spécialistes dans quelques grandes unités.

Le choix de la politique tarifaire de la télémedecine est également un problème important à résoudre.

Enfin, l'hétérogénéité des besoins de chaque praticien et des normes informatiques impose de développer des applications à un degré de compatibilité important.

A.4 Enjeux et nécessité

Le marché mondial de la télémedecine est un secteur économique à fort potentiel de développement. La santé devrait être amenée à représenter une bonne part du chiffre d'affaire mondial des télécommunications. Le prix des équipements en télémedecine a déjà beaucoup baissé.

La Norvège, l'Angleterre et les Etats-Unis sont particulièrement actifs dans le domaine de la télémedecine. La France, quant à elle, est confrontée à des difficultés pour la mise en place d'une politique de développement et de régularisation des projets de télémedecine [2]. Face au problème de désordre des normes informatiques et télématiques, à l'émergence de projets disparates, parfois concurrentiels, la nécessité d'un projet global pour le système de santé semblerait pourtant devoir s'imposer.

La télémedecine s'avère cependant être une réalité médicale : elle s'impose déjà à travers l'usage d'outils comme le téléphone, la télécopie par exemple. Les progrès actuels des NTIC (imagerie médicale, débits de transmission, convivialité des systèmes...) appliquées au domaine médical ouvrent des perspectives d'amélioration de la qualité des soins. Par ailleurs, la télémedecine devrait contribuer significativement à contenir l'explosion des dépenses de santé. Elle impose cependant beaucoup d'investissements et la formation du personnel médical ainsi que celle des patients dans certains cas (télésurveillance à domicile par exemple). Les changements de mentalité des professionnels de la santé et une meilleure gestion des réseaux de soins sont des facteurs-clés du succès de son développement.

Annexe B Description de l'appartement prototype

Pour les besoins de l'expérimentation, un appartement prototype de type F1 (50 m² environ) a été mis en place à la Faculté de médecine de Grenoble. Un certain nombre de capteurs en état de fonctionnement y sont déjà installés. Il pourra être envisagé l'installation d'autres types de capteurs que ceux déjà disponibles, par exemple des capteurs sonores dans l'objectif de détecter un appel de détresse.

L'ensemble de ces données est utilisée pour la détection des situations critiques de la personne à domicile ou la construction des courbes de tendance relatives au patient. On considère notamment les paramètres reçus des capteurs en temps réel, mais aussi l'évolution temporelle de ces valeurs ainsi que d'autres données connues du système.

B.1 Les capteurs

B.1.1 Capteurs de présence / mouvement

Les capteurs de présence sont des détecteurs de mouvement infra-rouge DP8111 – DAITEM. Cinq de ces capteurs sont répartis dans l'appartement prototype, définissant ainsi **5 zones de surveillance** :

- la *chambre*
- le *séjour*
- la *cuisine*
- la *salle de bain*
- le *couloir*

Ces capteurs sont autonomes et intègrent un émetteur radio. Ils permettent la surveillance de locaux par détection d'une brusque variation de rayonnement infra-rouge ou de température dans la zone surveillée. **Trois types de détection** sont possibles lorsque ces capteurs sont associés à une centrale de la gamme DP8511 :

- *détection volumétrique*, adaptée aux pièces principales du logement (portée de 12m)
- *détection linéaire*, adaptée aux pièces longues et étroites (portée de 15m)
- *détection rideaux* pour la surveillance des issues (portée de 10m)

La configuration des capteurs admet la **sémantique booléenne** suivante :

- 0 = *présence/ mouvement*
- 1 = *aucun mouvement*

B.1.2 Capteurs de contacts de porte

Ces capteurs sont des émetteurs universels DP8211 – DAITEM. Deux de ces capteurs sont installés dans l'appartement prototype, définissant ainsi **deux zones de surveillance** :

- la porte d'*entrée*
- la porte de la *salle de bain*

Ces capteurs sont autonomes et intègrent un émetteur radio. Ils permettent la détection de l'ouverture des portes.

La configuration de ces capteurs admet la **sémantique booléenne** suivante :

- 0 = *action d'ouverture de la porte*
- 1 = *aucune action d'ouverture détectée*

B.1.3 Tensiomètre

Le tensiomètre fournit des informations sur le rythme cardiaque de la personne. **Quatre valeurs numériques** sont transmises au système :

- la *Pression Systolique* (PS)
- la *Pression Diastolique* (PD)
- la *Pression Moyenne* (PM)
- le *Fréquence Cardiaque* (FC)

B.1.4 Pèse-personne

Le pèse-personne permet de connaître l'évolution du poids de la personne. Ce poids est transmis au système sous forme d'une valeur numérique.

B.1.5 Actimètre

L'actimètre est embarqué sur la personne et dispose d'une liaison radio avec le système. Il inclut dans un même boîtier quatre capteurs :

- deux capteurs de position de la personne : *verticale* et *horizontale*.
- un capteur d'accélération verticale (*accéléromètre*) permettant la détection des chutes.
- un capteur de vibration corporelle permettant de détecter de petits signes de vie sur la personne. Ce capteur permet notamment de connaître l'évolution du rythme cardiaque.

Le message transmis au système comporte quatre indicateurs ou descripteurs définis à partir des capteurs de l'actimètre :

- un *indicateur de position* (p0) :
 - 0 = *debout*
 - 1 = *allongée*
- un *indicateur de chute* (a0)
 - 0 = *chute*
 - 1 = *pas de chute*
- un *indicateur de vibration* (n0) :
 - 0 = *pas de vibrations détectées*
 - 1 = *vibrations détectées*
- un *indicateur d'urgence* (u0) : la valeur de cet indicateur est calculée par combinaison (laquelle ?) des valeurs des autres indicateurs.
 - 0 = *urgence*
 - 1 = *pas d'urgence*

B.2 Transmission des informations des capteurs

B.2.1 Identifiant du capteur

A chaque type de capteur est associée une *carte d'acquisition* permettant de mettre en forme le message à envoyer au système. Le protocole utilisé pour le formatage des informations est en effet spécifique à chaque type de capteur puisqu'il dépend des données à transmettre. Chaque message envoyé au système comporte un *identifiant* précisant de quel capteur provient l'information contenue dans le message.

Les capteurs volumétriques (ou de présence) et les capteurs de contact de porte sont reliés à la même carte. Les informations en provenance de ces capteurs sont donc intégrées dans un seul message envoyé au système. Ainsi, il n'y a que 4 valeurs d'identifiant possibles pour définir l'origine du message :

- *capteurs volumétriques (incluant les capteurs de contact de porte)*
- *tensiomètre*
- *pèse-personne*
- *actimètre*

B.2.2 Paramètres des messages

Le Tableau 2 présente un récapitulatif des messages pouvant être reçus par le système.

Tableau 2. Les messages reçus par le système et leurs paramètres

Type de message	Paramètres reçus	Valeurs possibles
Capteurs volumétriques	Présence dans le couloir	0/1
	Présence dans la salle de bain	0/1
	Présence dans la cuisine	0/1
	Présence dans le séjour	0/1
	Présence dans la chambre	0/1
	Ouverture de la porte d'entrée	0/1
	Ouverture de la porte de la salle de bain	0/1
Tensiomètre	Pression Systolique (PS)	Valeur numérique
	Pression Diastolique (PD)	Valeur numérique
	Pression Moyenne (PM)	Valeur numérique
	Fréquence Cardiaque (FC)	Valeur numérique
Pèse-personne	Poids	Valeur numérique
Actimètres	Position allongée / debout	0/1
	Chute	0/1
	Vibrations	0/1
	Urgence	0/1

B.2.3 Système central de l'HIS

Les messages envoyés par les capteurs sont récupérés et décodés par une carte CAN (« Controller Area Network ») au niveau du système central.

Les données sont stockées dans des fichiers texte sous un format de type XML. Chaque message reçu ajoute une ligne dans ces fichiers. La date et l'heure de réception sont toujours précisées (tags <date> et <heure>), suivies des données spécifiques à chaque capteur. Par exemple pour les capteurs de présence, les tags <entrée> et </entrée> encadrent le nom de l'endroit où un mouvement a été détecté.

Annexe C Le système cardiovasculaire

C.1 Description du système cardiovasculaire

Le système cardiovasculaire de l'homme comporte le cœur et deux réseaux parallèles de vaisseaux : les artères et les veines.

L'énergie nécessaire à l'organisme est apportée aux cellules par le réseau sanguin. Le **sang** est un liquide physiologique de couleur rouge. Il circule dans les vaisseaux et apporte aux tissus qu'il irrigue l'oxygène et les substances nutritives. Il transporte également vers les organes éliminateurs le gaz carbonique et les déchets du métabolisme.

Le **cœur** (Figure 23) assure la distribution du sang dans tout l'organisme. Sa principale fonction est de recevoir le sang sous faible pression et de le retourner dans les artères à une pression suffisamment élevée pour que celui-ci parcoure tous les vaisseaux et atteigne à nouveau le cœur. Le cœur un organe musculaire creux divisé en quatre cavités : les **oreillettes**, droite et gauche, et les **ventricules**, droit et gauche. Le muscle qui le compose est spécifique : c'est le **myocarde**. Les oreillettes reçoivent le sang des **veines**, tandis que les ventricules envoient le sang dans les **artères**. Les artères constituent donc le réseau des vaisseaux délivrant l'oxygène, les veines constituant le réseau « de retour » du sang au cœur et aux poumons. Les oreillettes sont petites et situées dans la partie supérieure. Les ventricules sont plus gros et situés sous les oreillettes. Un jeu de clapets, appelés **valves**, entre oreillettes et ventricules et entre ventricules et artères, assure l'écoulement en sens unique du sang.

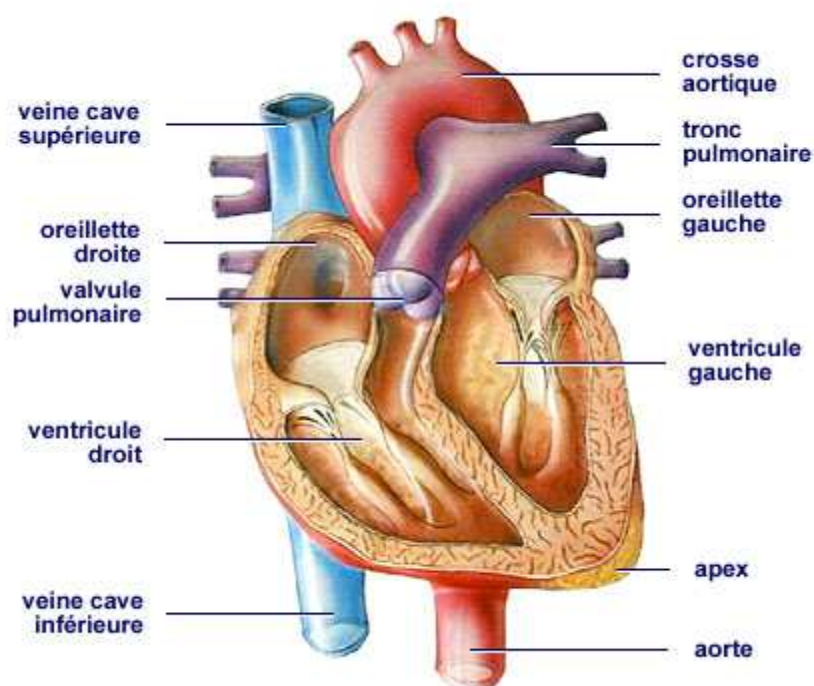


Figure 23. Schéma en coupe du cœur [47]

Le cœur est en fait un organe double : on distingue un **cœur droit** et un **cœur gauche**. Le cœur droit met en communication l'oreillette et le ventricule droit, chargés de propulser le sang désoxygéné jusqu'aux poumons par l'intermédiaire de l'artère pulmonaire, constituant ainsi la **petite circulation**. Le cœur gauche est quant à lui composé de l'oreillette et du ventricule gauche. Son rôle est de recueillir le sang oxygéné provenant des poumons et de le propulser dans tout l'organisme par l'intermédiaire de l'**aorte**, tronc d'origine de toutes les artères du corps : c'est la **grande circulation**.

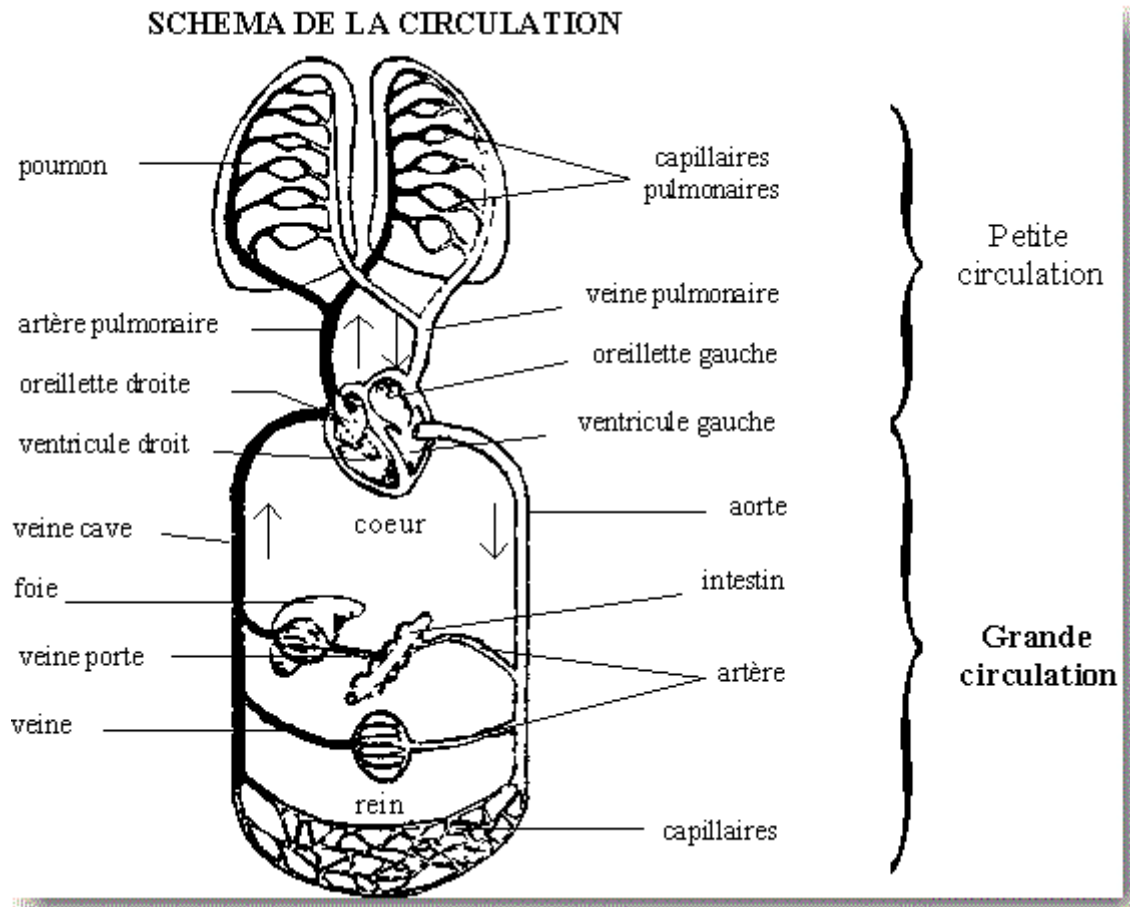


Figure 24. Schéma de la circulation générale et de la circulation pulmonaire [48]

Du point de vue fonctionnel, le cœur est donc une double pompe placée au point de raccordement des deux circulations (Figure 24) :

- Le cœur droit est la pompe de la **circulation pulmonaire**, destinée aux échanges gazeux. Le sang des **veines systémiques**, pauvre en oxygène, entre dans l'oreillette droite, passe dans le ventricule droit, emprunte le tronc pulmonaire pour se rendre aux poumons et revient oxygéné, dans l'oreillette gauche par les veines pulmonaires.
- Le cœur gauche est la pompe de la **circulation systémique**. Le sang, riche en oxygène, provenant des poumons, entre dans l'oreillette gauche, s'écoule dans le ventricule gauche et emprunte l'aorte, dont les ramifications le distribuent dans tout l'organisme. Les veines systémiques ramènent le sang pauvre en oxygène dans l'oreillette droite.

Chaque cycle de l'activité cardiaque chasse le sang des ventricules pour l'envoyer dans les circulations pulmonaires et systémiques. Un cycle cardiaque s'effectue en deux phases : la contraction du cœur, ou *systole*, et la phase de relâchement, ou *diastole*.

- La **systole** est la phase de contraction qui expulse le sang des chambres ou cavités du cœur vers le reste de l'organisme. Les oreillettes et les ventricules se contractent l'un après l'autre. On peut ainsi distinguer deux phases de contraction : la contraction des oreillettes (systole auriculaire) et la contraction des ventricules (systole ventriculaire).

La **systole auriculaire** va propulser dans les deux ventricules durant leur relâchement simultané (diastoles ventriculaires) le sang contenu dans les oreillettes droite et gauche.

La **systole ventriculaire** envoie le sang dans les artères, aorte à gauche et pulmonaire à droite, pendant le relâchement des oreillettes (diastoles auriculaires). Le sang chassé par l'artère pulmonaire (ventricule droit) passe dans la petite circulation et simultanément chassé par l'aorte (ventricule gauche) va passer dans la grande circulation.

- La **diastole générale** est la phase de relâchement simultané des oreillettes et des ventricules qui permet le remplissage sanguin des oreillettes.

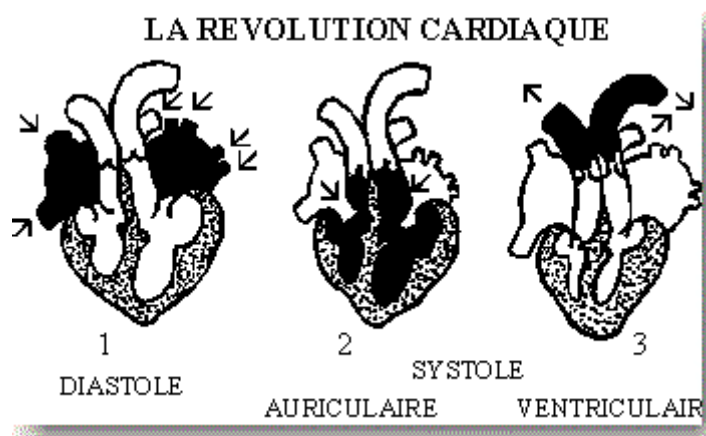


Figure 25. Schéma général de la révolution cardiaque [48]

La **Révolution Cardiaque (RC)** (Figure 25) est la suite des événements qui se produisent pendant un **battement**, c'est-à-dire entre une contraction du cœur de la suivante. Chaque révolution cardiaque comprend une contraction et un relâchement complet des oreillettes et des ventricules.

La **fréquence** ou **rythme cardiaque** mesure le nombre de RC par minute. Au repos, les valeurs sont proches de 60 à 70 cycles par minute. La fréquence cardiaque varie selon plusieurs facteurs :

- L'âge : la fréquence cardiaque diminue avec l'âge.
- L'entraînement : le sportif voit sa fréquence de repos et d'effort baisser par rapport au sujet sédentaire pour le même effort.
- La fréquence cardiaque augmente du fait de la thermorégulation (fièvre), lors de la digestion, au changement de la position orthostatique (debout) par rapport à la position décubitus allongée (position allongée du corps en état de repos), chez le fumeur sous l'influence de la nicotine, chez les consommateurs de stimulants (caféine, théine, amphétamines), etc.

Le **débit cardiaque** est le volume de sang expulsé par chaque ventricule au cours de l'unité de temps. Deux facteurs le déterminent : le volume éjecté au cours de chaque contraction et la fréquence à laquelle le cœur se contracte. Dans les conditions de repos, le débit cardiaque est d'environ 5 litres par minute chez un adulte moyen dont la fréquence cardiaque est de l'ordre de 70 contractions par minute. Le débit cardiaque s'adapte aux besoins de l'organisme : il varie par exemple avec la chaleur (vasodilatation cutanée : augmentation du calibre des vaisseaux sanguins relatifs à la peau), l'altitude (réduction du retour du sang au cœur droit), les émotions, l'absorption d'aliments. Mais c'est surtout l'exercice physique qui peut l'élever, essentiellement par une augmentation de la fréquence cardiaque.

C.2 La pression artérielle

C.2.1 Définitions

L'existence de la pression artérielle a pour but de pousser le sang et les globules rouges jusqu'au niveau des tissus de l'organisme pour qu'ils puissent assurer leur fonction : oxygéner les tissus. Au niveau de ces tissus, la pression artérielle est très faible et la vitesse très ralentie de manière à ce que l'oxygène puisse mieux pénétrer dans les tissus.

La **pression artérielle** est la pression exercée par le flux de sang sur la paroi des artères. Elle se définit comme le produit entre la résistance du vaisseau dans lequel se trouve le sang et le débit imposé par le cœur.

$$\text{Pression artérielle} = \text{Résistance du vaisseau} \times \text{Débit cardiaque}$$

La **résistance du vaisseau** correspond à la rigidité des parois de l'artère, qui peut varier dans de très nombreuses circonstances.

Le **débit cardiaque** correspond à la capacité du cœur à éjecter du sang lors de chaque contraction (ou systole), correspondant donc à un volume de sang éjecté par minute (environ 5 litres par minute en moyenne).

La pression artérielle résulte de la pression exercée par le cœur et varie naturellement au rythme du cœur. Un cycle cardiaque est réalisé en deux grandes étapes : la contraction du cœur, ou **systole**, et la phase de relâchement, ou **diastole**. Au cours de chaque révolution cardiaque, la pression artérielle oscille entre une valeur maximale ou **pression artérielle systolique (PS)** et une valeur minimale ou **pression artérielle diastolique (PD)**.

- **Pression artérielle systolique (PS)**

Lors de la phase de contraction du cœur (systole), le ventricule gauche se contracte et expulse le sang dans l'aorte (systole ventriculaire). Le sang étire alors les parois élastiques de l'aorte et la pression devient maximale. La pression artérielle systolique ou maximale correspond donc au moment où le cœur a fini sa contraction.

- **Pression artérielle diastolique (PD)**

Les valves de l'aorte se referment ensuite pour que le sang ne reflue pas (diastole ventriculaire). Les parois des artères reviennent alors en position initiale et la pression devient minimale. La pression artérielle diastolique ou minimale correspond donc à la fin de la décontraction du cœur.

La différence de pression est nécessaire à l'écoulement du sang de la sortie de l'aorte (pression systolique) vers les organes (pression diastolique). La **tension artérielle** s'exprime par ces deux valeurs : la pression maximale (systolique) et la pression maximale (diastolique). Elle est exprimée en millimètres de mercure (mm Hg), 140/85 par exemple, ou plus usuellement en centimètres de mercure : 14/8.5.

Le schéma de la Figure 26 présente l'allure de la courbe d'évolution de la pression au cours d'une révolution cardiaque.

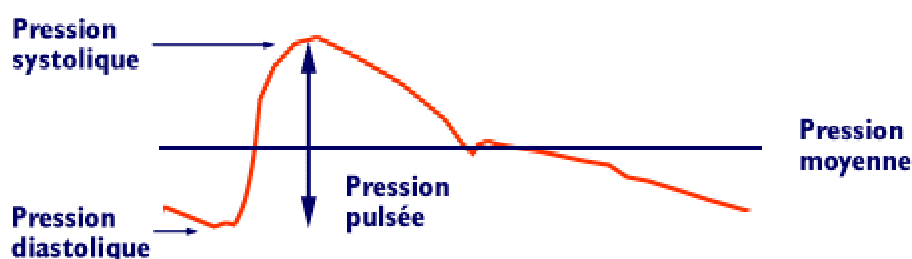


Figure 26. Enregistrement de la pression artérielle à l'aide d'un capteur introduit dans l'aorte [47]

La **pression artérielle différentielle** est la différence entre la pression systolique et la pression diastolique.

La **pression artérielle moyenne** se définit par la surface sous la courbe de pression lors d'un cycle cardiaque divisée par la durée du cycle. Elle peut être estimée à partir des pressions diastolique et différentielle en ajoutant à la pression diastolique le tiers de la pression différentielle.

A chaque révolution cardiaque, l'alternance de la dilatation et du resserrement des artères élastiques résultant de la contraction du cœur crée une onde de pression (sensation de soulèvements intermittents). L'ondée sanguine prend son origine lors de la systole ventriculaire dans l'aorte.

Le **pouls** est le battement d'une artère superficielle correspondant à la transmission par le sang de cette onde pulsatile déclenchée par chaque contraction cardiaque. Il indique donc le rythme cardiaque. Le rythme cardiaque normal est en moyenne de 70 cycles par minutes, sauf chez les sportifs pour lesquels il est d'environ 40 à 50 pulsations par minute. Sauf cas particulier, un pouls inférieur à 50 est trop bas. On estime que la fréquence cardiaque doit être comprise entre 50 et 90 pulsations par minute. La fréquence maximale théorique est de 220 moins l'âge (soit par exemple 170 à 50 ans).

C.2.2 Fluctuations de la pression artérielle

La pression artérielle varie en fonction de nombreux critères pouvant être classés en deux grandes catégories : les critères de variation selon la localisation de l'artère considérée par rapport au cœur et selon le temps.

- **Selon la localisation**

Plus le sang s'éloigne du cœur par les artères et moins la pression artérielle sera importante car l'oxygène transporté par le sang doit avoir le temps de passer dans les organes.

En pratique, il a été constaté que la pression artérielle mesurée au niveau du bras est proche de celle mesurée dans l'aorte. Mais la pression artérielle mesurée au niveau du poignet semble moins fiable.

- **Selon le temps**

La pression artérielle fluctue à court, moyen ou long terme.

Il existe un cycle de pression artérielle **sur la durée de la vie** d'un individu. En effet, la pression artérielle est plus élevée chez les sujets âgés que chez les sujets jeunes. L'âge est ainsi un facteur important de variation.

La pression artérielle varie également selon les saisons, puisqu'elle est plus basse en été qu'en hiver (Figure 27).

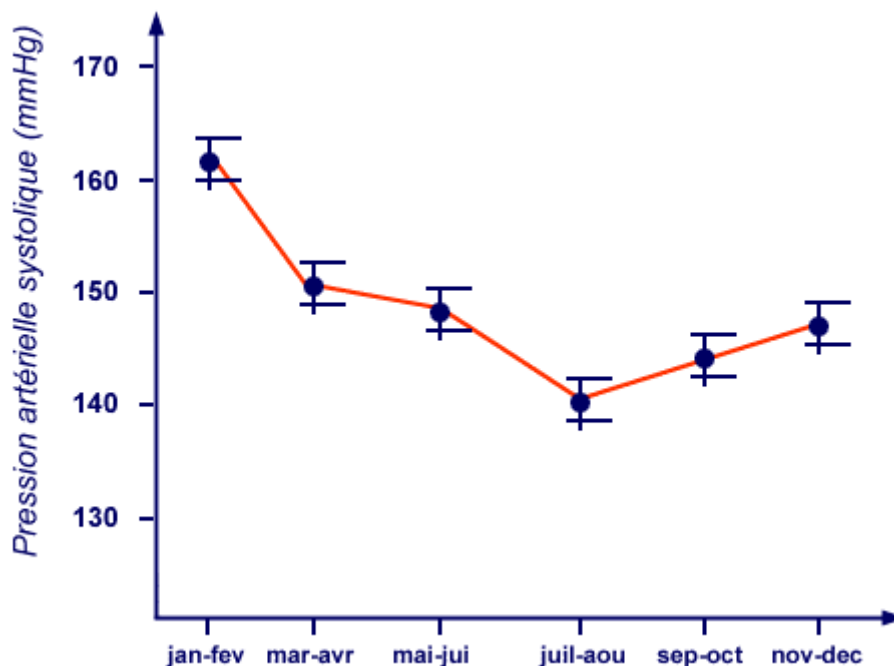


Figure 27. Variations de la pression artérielle systolique durant l'année [47]

La pression artérielle est aussi une valeur clinique qui varie au cours d'une journée (Figure 28). L'activité cardiaque est régulée afin de permettre une réponse adaptative du cœur aux besoins de l'organisme, par exemple en cas d'effort musculaire. La pression artérielle est plus élevée en période d'activité, c'est à dire le jour, qu'en période de sommeil. On la mesure habituellement au repos mais on peut également l'enregistrer au cours d'une journée en ambulatoire.

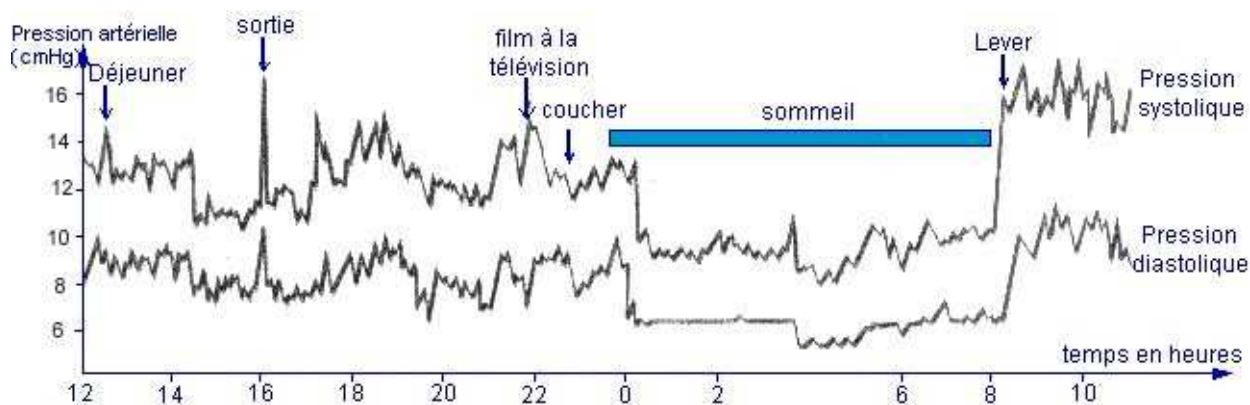


Figure 28. Variation de la pression artérielle au cours d'une journée fournies par un enregistrement de la pression artérielle grâce à un appareil de mesure ambulatoire [49]

Il est important de noter la variation de la pression artérielle entre la période de jour et celle de nuit. Normalement, la pression artérielle est inférieure durant la période de nuit par rapport à la période de jour et cette variation doit être au moins égale à 10%. Lorsque ce cycle n'est pas respecté, dans le cas d'un sujet hypertendu par exemple, cela témoigne de la sévérité de l'hypertension artérielle ou de l'existence d'une pathologie responsable d'une hypertension.

Enfin, la pression artérielle fluctue à bien plus court terme en fonction de nos émotions, de la fréquence respiratoire et de nos commandes nerveuses qui la régulent automatiquement.

Les **facteurs de variation** de la pression artérielle à court et moyen terme sont la position, le froid, le stress, l'effort par exemple.

La pression artérielle augmente avec l'effort, c'est pourquoi elle doit être mesurée allongée au repos. En effet, lors d'un intense exercice physique, les muscles ont besoin de plus d'éléments nutritifs ce qui provoque une augmentation du débit sanguin, donc une variation positive de la pression.

La variation de pression peut être également due aux émotions. Les vaisseaux sanguins se dilatent, le débit est plus fort donc la pression systolique augmente (le visage devient alors tout rouge par exemple). La pression artérielle varie avec le stress (tension nerveuse). Les hormones peuvent également faire varier le débit sanguin : l'adrénaline par exemple augmente la force de contraction du cœur ainsi que le nombre de battements pour de courts moments.

La variation de pression est caractéristique suivant l'individu.

Un sportif, du fait de son entraînement, aura une pression plus faible qu'un cardiaque souffrant de pathologies vasculaires.

Certaines pathologies induisent une augmentation de la pression artérielle (insuffisance rénale, athérosclérose par exemple) alors que d'autres entraînent sa diminution (hémorragie) produisant une hypotension. Lors d'une hémorragie, le volume de sang devient plus faible. Le cœur accélère son rythme de battements pour compenser le manque de liquide rouge tandis que la pression diminue fatalement.

C.2.3 Principe de mesure de pression artérielle

La méthode la plus pratiquée est la mesure de la pression artérielle à l'aide d'un brassard. Il s'agit d'amener d'abord la pression du brassard (PB) à une valeur supérieure à la pression systolique pour bloquer la circulation artérielle dans le bras. On laisse ensuite la pression du brassard diminuer progressivement jusqu'à la valeur limite à partir de laquelle la pression artérielle est suffisante pour laisser de nouveau passer le sang dans l'artère : c'est la pression systolique (PS). En poursuivant le dégonflage, on amène la pression du brassard à une valeur à partir de laquelle il n'y a plus d'obstacle au flux artériel même lorsque le cœur est en diastole (décontraction) : c'est la pression diastolique (PD).

Le stéthoscope permet d'identifier les deux moments clés de la mesure en écoutant le pouls de l'artère brachiale. Lorsque le brassard est gonflé au dessus de la pression systolique on n'entend rien puisque le sang ne passe pas. En dégonflant lentement le brassard, on commence à détecter le pouls artériel dès que l'on passe en dessous de la pression systolique car désormais le sang recommence à circuler et se fait entendre à chaque systole. Le brassard continuant à se dégonfler, on n'entend plus le pouls artériel dès que l'on passe en dessous de la pression diastolique.

Le schéma de la Figure 29 montre la relation entre pression artérielle, pression dans le brassard et bruits de l'artère.

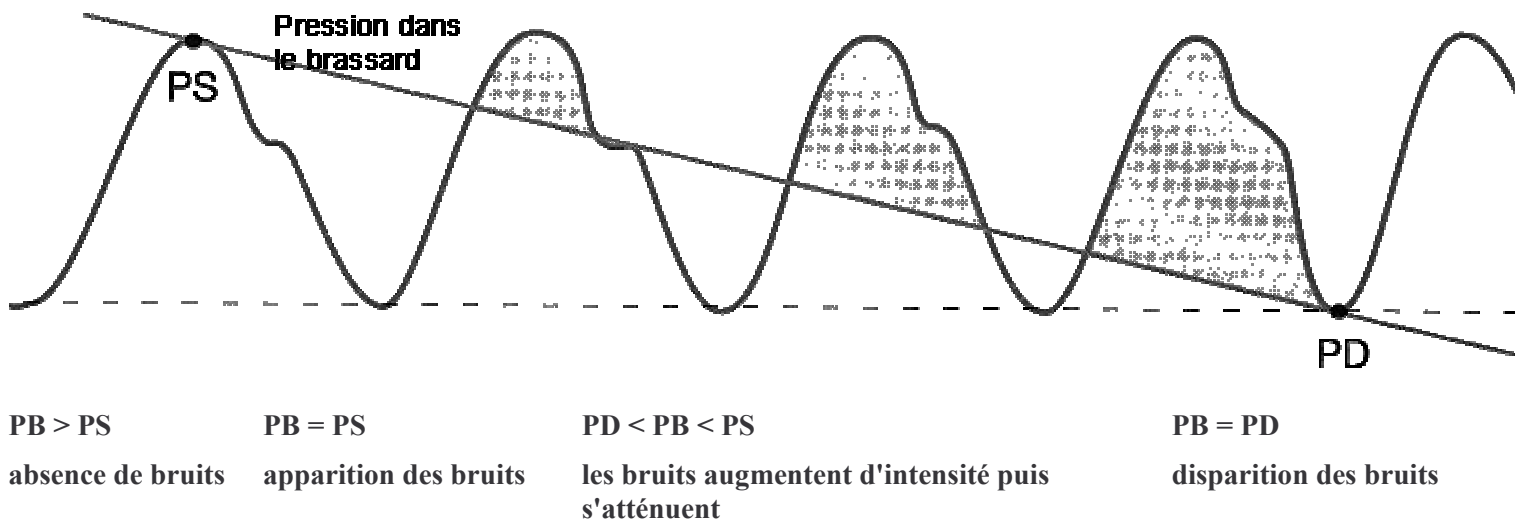
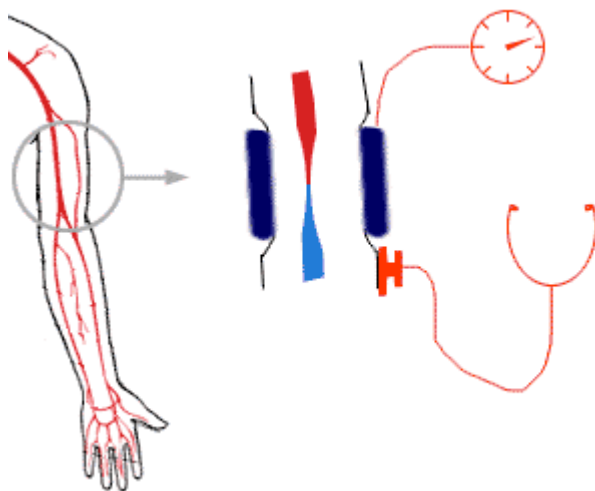


Figure 29. Relation entre l'évolution de la pression artérielle (cycle cardiaque), la pression du brassard et le bruit de l'artère [50]

PB : Pression du Brassard – PS : Pression Systolique – PD : Pression Diastolique

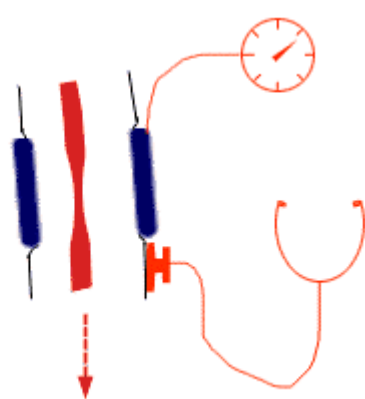
Les schémas de la Figure 30 décrivent la manipulation correspondant à la prise de la tension artérielle.



– 1 –

Le matériel de mesure comprend un brassard gonflable, un cadran de mesure et un stéthoscope.

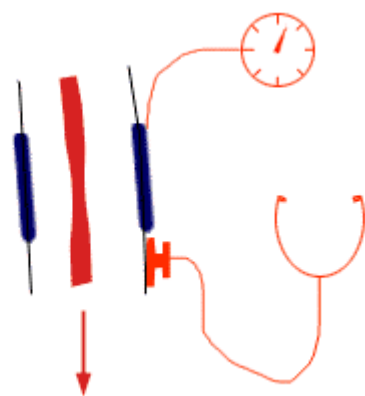
Le brassard comprime l'artère du bras, le sang ne passe plus : aucun bruit n'est perçu par le stéthoscope.



– 2 –

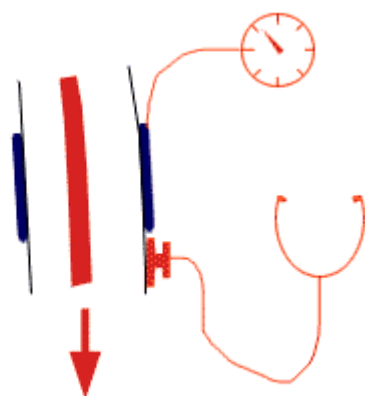
Puis le brassard est dégonflé lentement, le sang commence à passer dans l'artère : un bruit est perçu par le stéthoscope et la valeur de la pression artérielle est lue au même moment sur le cadran.

La pression artérielle mesurée à cet instant est la pression artérielle maximale, c'est-à-dire la pression artérielle **systolique**.



– 3 –

Le brassard continue de se dégonfler. Le sang passe de mieux en mieux et un bruit est toujours perçu par le stéthoscope.



– 4 –

Plus le brassard se dégonfle et moins le bruit est audible par le stéthoscope, jusqu'au moment où il disparaît : la pression artérielle est alors lue sur le cadran et définit la minima, c'est à dire la pression artérielle **diastolique**.

Figure 30. Les étapes de la prise de tension [47]

C.2.4 Pratique de la mesure

Pour être convenablement réalisée, la mesure de pression artérielle doit se faire en dehors de tout effort ou de période de stress, dans un endroit calme, et après un repos d'environ 5 minutes. La prise de la tension artérielle doit être effectuée dans différentes positions (allongée et debout) et aux deux bras. Le bras dominant (le bras droit pour un droitier ou gauche pour un gaucher) possède une pression artérielle plus élevée que l'autre bras. Il est nécessaire d'autre part que le bras supportant l'appareil de tension artérielle soit le plus relâché possible.

On mesure généralement la pression dans l'artère brachiale. Pour cela, le brassard est placé autour du bras et le stéthoscope sur la saignée du coude à l'emplacement où le pouls est détectable avec le bout des doigts.

On considère que la tension normale chez un adulte ne doit pas dépasser 14/9, ou encore 140/90 mm Hg, au repos lorsqu'elle est mesurée en consultation. La tension est généralement un peu plus basse

lorsqu'elle est mesurée par la personne elle-même, en-dehors du cabinet médical, à l'aide d'un appareil d'auto-mesure. Les valeurs normales de tension sont alors de 135/85 mm Hg.

C.3 Les risques cardiovasculaires

C.3.1 Hypertension

L'hypertension artérielle (HTA) est définie comme une élévation de la pression du sang dans les artères par rapport à une valeur dite "normale" établie par de nombreux comités scientifiques à travers le monde. On parle d'hypertension artérielle quand la tension artérielle est **supérieure à 160 / 95 mm de mercure**, de façon permanente ou pas, quand un individu est au repos.

Ces critères sont directement en relation avec l'âge du patient. La règle de calcul pour une pression artérielle systolique normale (pression maximale) est la suivante : $PS = \text{âge de l'individu} + 100$.

Cela signifie qu'un individu âgé de 20 ans doit avoir 120 mm de mercure pour le chiffre correspondant à la pression maximale (pression artérielle systolique), soit 12 cm de mercure (expression d'usage). Elle sera de 13 pour un individu de 30 ans, de 14 pour un individu de 40 ans, ainsi de suite.

La pression artérielle diastolique (deuxième chiffre) est considéré comme anormale quand elle dépasse 95 mm de mercure (exprimé usuellement par 9,5 cm de mercure). Si on considère uniquement ce chiffre, l'hypertension est dite :

- **Légère** entre 95 et 105
- **Modérée** entre 105 et 115
- **Grave** au-dessus de 115 millimètres

En France, environ 10% de la population générale est porteuse d'une hypertension artérielle, soit environ 6 millions de personnes.

L'hypertension peut être liée à un autre problème médical : atteinte des artères, pathologie des reins ou des glandes surrénales (deux petites glandes situées au-dessus des reins) par exemple. On parle alors d'**hypertension secondaire**. Ces situations sont cependant relativement rares.

Dans l'immense majorité des cas (95%), l'hypertension est dite « **essentielle** », c'est-à-dire que la cause en est inconnue. On en connaît cependant quelques symptômes :

- Céphalées (maux de tête) de localisation occipitale (en arrière du crâne) surtout pendant la deuxième moitié de la nuit ou au réveil.
- Fatigue
- Perte de mémoire
- Troubles visuels (perte transitoire de la vue, éblouissement)
- Vertiges
- Bourdonnements d'oreille
- Troubles visuels
- Epistaxis (saignement de nez)
- Dyspnée (essoufflement)
- Polyurie (sécrétion d'urine en quantité abondante)
- Pollakiurie (fréquence exagérée des mictions : expulsion des urines accumulées dans la vessie)
- Palpitations

S'il n'existe aucune cause évidente de l'hypertension chez la plupart des malades, il y a en revanche une multitude de facteurs susceptibles de favoriser la survenue d'une hypertension artérielle.

- **L'hérédité et l'âge** : ce sont des facteurs contre lesquels on ne peut pas lutter. Lorsque l'on a des parents hypertendus (le père, la mère ou a fortiori les deux) le risque de développer soi-même une hypertension est augmenté. De même, le risque de développer une hypertension augmente avec l'âge. Ainsi, il faut tenir compte de l'âge du patient dans l'évaluation du risque cardiovasculaire global. Une personne de 65 ans qui présente une hypertension artérielle à 170/95 millimètres de mercure à beaucoup plus de risque qu'une personne âgée de 45 ans.
- **Le sexe** : en règle générale, les hommes sont plus concernés par l'hypertension artérielle que les femmes non ménopausées. La différence est de l'ordre de 6 à 7 millimètres de mercure pour la pression artérielle maximale (systolique) et de 3 à 5 millimètres de mercure pour la pression artérielle minimale (diastolique). Néanmoins, cette différence s'annule à partir du moment où les femmes sont ménopausées.
- **Le mode de vie** : le surpoids ou l'obésité, une alimentation trop salée, l'abus d'alcool, de réglisse sont connus pour faciliter le développement d'une hypertension et l'entretenir lorsqu'elle existe.
- **Certains médicaments** également comme la cortisone, la contraception orale augmentent les risques d'hypertension.
- **Le stress** entretient l'hypertension plus qu'il n'en est la cause.
- **Le fait d'être de race noire** : les personnes de peau noire sont plus fréquemment hypertendues que les sujets à peau blanche et leur hypertension est plus souvent résistante au traitement.

L'hypertension est un facteur de risque de **pathologie cardiovasculaire** (relatif au cœur et aux vaisseaux sanguins). Mais il en existe bien d'autres. Plus on cumule les facteurs de risque, plus le risque de développer une pathologie cardiovasculaire augmente.

Il est donc important de rechercher puis de contrôler chaque fois que cela est possible l'ensemble des facteurs de risque présents. La plupart peuvent être modifiés ou maîtrisés alors qu'il y en a certains sur lesquels on ne peut pas agir.

Les principaux facteurs de risque de pathologie cardiovasculaire sont :

- **Le sexe masculin** : les femmes sont relativement protégées jusqu'à la ménopause.
- **L'hérédité** : une histoire familiale de pathologie cardiaque précoce avant l'âge de 65 ans chez votre mère ou une sœur, ou avant l'âge de 55 ans chez votre père ou un frère augmente votre propre risque d'accident cardiaque.
- **L'âge** : au delà de 45 ans pour les hommes et de 55 ans pour les femmes.
- **Le tabagisme** : la consommation de tabac est responsable d'une augmentation des accidents cardiovasculaires.
- **Un taux élevé de cholestérol** dans le sang augmente le risque de présenter une complication cardiovasculaire, que ce soit un infarctus ou un accident vasculaire cérébral.
- **L'hypertension artérielle.**
- **Le diabète** : il correspond à la présence d'un taux élevé de glucose dans le sang (glucose > 1,26 gramme/litre). L'existence d'un diabète isolé majore le risque cardiovasculaire, mais ce risque est encore plus grand lorsque le diabète est associé à une hypertension artérielle.
- **Le surpoids et a fortiori l'obésité.**
- **Un mode de vie** sédentaire.

Il est possible d'évaluer la gravité du risque cardiovasculaire global en fonction de la pression artérielle habituelle d'un individu et du nombre de facteurs de risque cardiovasculaire supplémentaires qui le concernent (Tableau 3).

Tableau 3. Gravité du risque cardiovasculaire en fonction des facteurs de risques du patient [47]

Autres facteurs de risque et antécédents pathologiques	Pression artérielle (mm Hg)		
	HTA légère PAS 140-159 ou PAD 95-104	HTA modérée PAS 160-179 ou PAD 105-114	HTA sévère PAS > 180 ou PAD > 115
I. Pas d'autres facteurs de risque	Risque faible	Risque moyen	Risque élevé
II. 1 à 2 facteurs de risque	Risque moyen	Risque moyen	Risque très élevé
III. 3 facteurs de risque ou plus, ou atteinte d'un organe	Risque élevé	Risque élevé	Risque très élevé
IV. Pathologies associées	Risque très élevé	Risque très élevé	Risque très élevé

Chez un sujet hypertendu, le principal risque est l'apparition d'une "crise aiguë hypertensive" (hypertension artérielle très sévère), correspondant à l'apparition brutale d'une hypertension artérielle très élevée : la pression artérielle maximale (systolique) est souvent supérieure à 250 millimètres de mercure (25) alors que la pression artérielle minimale (diastolique) est supérieure à 130 millimètres de mercure (13).

Le patient hypertendu peut également faire face à des poussées hypertensives. La pression artérielle est alors moins élevée qu'au cours d'une hypertension artérielle très sévère. Une telle poussée peut se manifester par un saignement de nez, des vertiges discrets, des maux de tête tolérables, une sensation d'oppression dans la poitrine.

Le diagnostic de l'hypertension s'effectue par la prise de la tension artérielle. Selon les caractéristiques du patient, les valeurs de pressions définissant une hypertension peuvent varier légèrement.

Pour les personnes diabétiques et/ou en insuffisance rénale par exemple, la valeur de pression artérielle à partir de laquelle est définie une hypertension artérielle est de 130/85 mm Hg.

Pour les sujets de plus de 60 ans, le chiffre de 160/90 mm Hg est toléré, en sachant que le chiffre de 140/90 mm Hg est idéal.

La mesure de la pression artérielle reste essentielle pour le suivi d'une hypertension artérielle.

C.3.2 Hypotension

L'**hypotension artérielle** est une diminution anormale de la pression à l'intérieur des artères ou de la cavité crânienne. Chez un adulte, on estime qu'il y a hypotension lorsque la pression artérielle systolique (pression artérielle maximale) est au-dessous de **100 mm de mercure**. Elle s'observe à l'état chronique chez les sujets fatigués, convalescents ou surmenés, ainsi qu'au cours de certaines affections chroniques, comme l'insuffisance surrénale (pathologie d'Addison). Pour un sujet à pression artérielle normale, toute chute soudaine à un chiffre moindre peut traduire un état de choc, une défaillance cardiaque, une hémorragie interne, une anémie ou un manque de tonicité des parois des vaisseaux sanguins par exemple. Une personne hypotendue peut avoir, notamment lorsqu'elle se lève brusquement ou après une station debout prolongée, des vertiges, un voile noir devant les yeux, des troubles du rythme cardiaque.

L'**hypotension orthostatique** est un symptôme caractérisé par la chute de la tension artérielle (pressions artérielles systolique et diastolique) lors du passage brusque de la position couchée à la position debout.

C.3.3 Arythmie cardiaque

L'étude du pouls permet une estimation rapide de l'état cardio-vasculaire d'une personne, et notamment la détection d'arythmies. L'**arythmie cardiaque** correspond à des perturbations du rythme cardiaque touchant sa fréquence, sa régularité et l'intensité de ses contractions. Ce trouble peut être d'origine physiologique ou faire suite à une pathologie. Il existe différentes variétés d'arythmie. On distingue notamment :

- Les **bradycardies** : diminution brutale et passagère du rythme cardiaque.
- Les **tachycardies** : accélération du rythme cardiaque.

C.4 Caractérisation des valeurs des mesures cardiaques

C.4.1 Pression Artérielle Systolique (PS)

La pression artérielle systolique est la pression artérielle maximale. Elle correspond au plus grand des deux chiffres représentant la tension artérielle. Sa valeur permet de diagnostiquer éventuellement une hypertension ou une hypotension.

L'ensemble des valeurs possibles de la pression systolique et leur signification sont représentés sur la Figure 31.

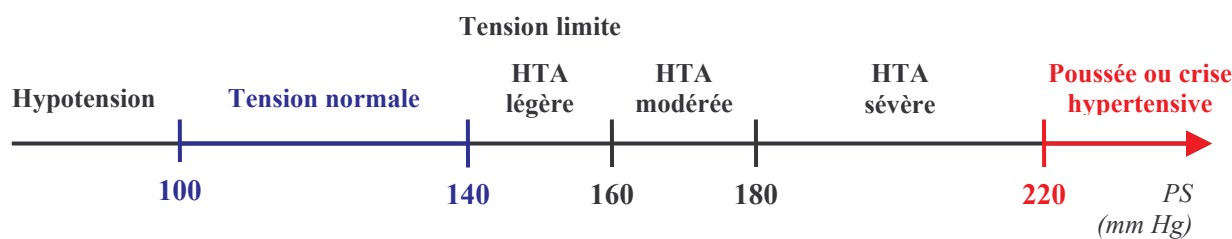


Figure 31. Echelle des valeurs de la pression artérielle systolique (PS)

C.4.2 Pression Artérielle Diastolique (PD)

La pression artérielle diastolique est la pression artérielle minimale. Elle correspond au plus petit des deux chiffres représentant la tension artérielle. Sa valeur permet de diagnostiquer éventuellement une hypertension.

L'ensemble des valeurs possibles de la pression diastolique et leur signification sont représentées sur la Figure 32.

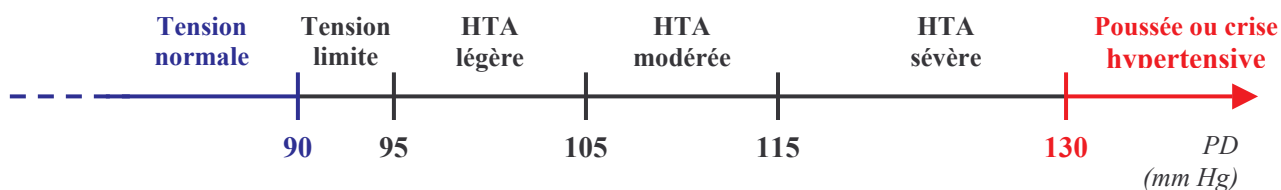


Figure 32. Echelle des valeurs de la pression artérielle diastolique (PD)

C.4.3 Fréquence Cardiaque

Le pouls correspond au battement d'une artère superficielle indiquant le rythme cardiaque.

L'ensemble des valeurs possibles du pouls à l'état de repos et leur signification sont représentées sur la Figure 33.

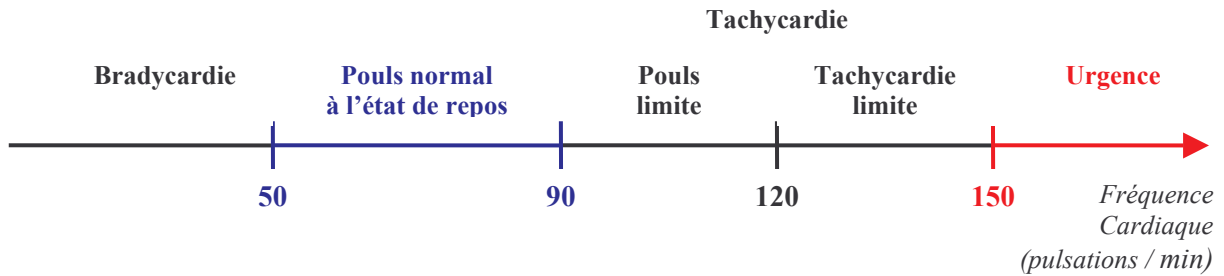


Figure 33. Echelle des valeurs de la fréquence cardiaque (FC)

Annexe D L'Indice de Masse Corporelle (IMC)

D.1 Définition et calcul

L'Indice de Masse Corporelle (IMC) est une mesure qui relie de façon mathématique la taille et le poids d'un individu. Il est encore appelé **indice de corpulence** ou **indice de Quételet**, du nom du statisticien et sociologue belge qui, dès le milieu du XIXe siècle, avait observé que cet indice était constant chez des individus de constitution normale [51].

Le calcul de l'IMC permet d'estimer la corpulence d'un individu et le risque éventuel lié à ce facteur pour sa santé. Cet indice s'obtient en divisant le poids (en kilogrammes) par le carré de la taille (en mètres). Il s'exprime donc en kg/m².

$$\text{IMC} = (\text{poids en kg}) / (\text{taille en m})^2$$

Par exemple, pour un poids de 60 kg et une taille de 1m60, l'IMC est de :

$$\text{IMC} = 60 / 1,6^2 = 60 / 2,56 = 23,43 \text{ kg/m}^2$$

D.2 Estimation de la corpulence d'un individu

Tout au long du XXe siècle, l'indice de Quételet a largement été utilisé pour étudier les proportions et la composition corporelles. Mais c'est seulement dans les années 1980 que son utilisation pour **mesurer le degré de l'obésité** a été proposée. Selon la classification acceptée au niveau international en 1998, le surpoids est défini par un IMC supérieur à 25,0 kg/m² et l'obésité par un IMC supérieur à 30,0 kg/m² avec une gradation en trois classes.

Ces définitions s'appliquent aux deux sexes et à tous les âges adultes jusqu'à 65 ans. Au-delà, et pour des groupes particuliers tels que les athlètes très entraînés, les enfants ou les femmes enceintes, ces seuils ne doivent pas être utilisés.

Des seuils identiques ont été retenus pour les hommes et pour les femmes. En effet, les valeurs retenues reposent essentiellement sur la mortalité associée à l'augmentation de l'IMC, et il se trouve que la mortalité relative associée à l'obésité est équivalente pour les deux sexes.

Au début des années 1990, l'IMC a également été proposé pour **estimer la maigreur** et le risque de déficit chronique en énergie chez l'adulte. Une classification à 3 seuils, 18,5 - 17,0 - 16,0, a été acceptée. Pour un individu, l'IMC est ainsi considéré comme normal lorsqu'il se situe entre 18,5 et 25.

D.3 Caractérisation de la valeur de l'IMC

La valeur de l'IMC permet de déterminer la corpulence de la personne, de voir s'il y a obésité ou maigreur, et d'en déterminer la sévérité (Tableau 4).

Tableau 4. Qualification des valeurs de l'Indice de Masse Corporelle (IMC)

IMC		Evaluation de la corpulence	
IMC < 18,5		Maigreur	
18,5 ≤ IMC < 25		Poids « normal »	
25 ≤ IMC < 30		Surpoids	
IMC ≥ 30	30 ≤ IMC < 35	Obésité	Obésité modérée
	35 ≤ IMC < 40		Obésité sévère
	IMC ≥ 40		Obésité massive ou morbide

Les médecins estiment que le risque pour la santé devient important au delà d'un indice de masse corporelle de 30 kg/m².

L'ensemble des valeurs possibles de l'IMC et leur signification sont représentées sur la Figure 34.

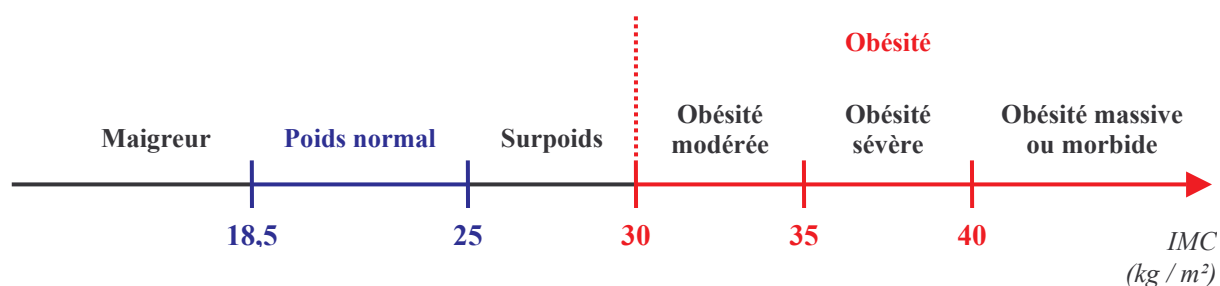


Figure 34. Echelle des valeurs de l'indice de masse corporelle (IMC)

Annexe E Troubles de la miction

E.1 Description

La miction est le mécanisme réflexe par lequel l'urine stockée dans la vessie est libérée à l'extérieur par le canal de l'urètre. L'urine est éliminée à raison de 5 à 6 mictions quotidiennes, en moyenne : habituellement de 0 à 1 pendant la nuit et de 4 à 5 environ dans la journée.

Le volume d'urine émis en 24 heures varie entre 60 ml chez le nouveau-né et 1,5 l en moyenne chez l'adulte. Il dépend directement de la quantité de liquide ainsi que de la teneur en eau des aliments ingérés. Chez un individu en bonne santé, la quantité normale d'urine excrétée quotidiennement varie entre ½ L et 2 L.

La **pollakiurie** est la fréquence excessive des mictions. Ces mictions ont lieu à intervalles fréquents et sont liées à une sensation de plénitude vésicale (impression de vessie pleine) qui n'est pas due à une vessie pleine mais à une vessie présentant une irritation. Cette sensation s'accompagne d'une impression de plénitude même quand la vessie n'est pas pleine. Il faut distinguer la **pollakiurie diurne** (pendant la journée) et la **pollakiurie nocturne** ou **permanente**.

La **nycturie** correspond à une excrétion urinaire ayant lieu préférentiellement la nuit. Normalement, la quantité d'urine diminue pendant la nuit (de 2 à 4 fois par rapport au jour). La nycturie se caractérise par une émission d'urine plus abondante la nuit que le jour. Elle force le malade à se lever une ou plusieurs fois pendant la nuit pour uriner, et se rencontre au cours du cancer de la vessie et de l'insuffisance rénale (le rein est incapable de concentrer les urines pendant la nuit). Elle est parfois due à l'absorption de boissons trop abondantes avant de se coucher. Elle correspond à une inversion du rythme normal de la diurèse (volume d'urine sécrétée par les reins pendant un certain temps).

La pollakiurie et la nycturie ne sont pas associées à une augmentation du volume total des urines. En revanche, la **polyurie** correspond à une augmentation des quantités d'urine émises par 24 heures (plus de 2 litres), et l'**oligurie** une diminution de la quantité d'urine émise pendant un laps de temps donné. Habituellement, on parle d'oligurie lorsque la diurèse (sécrétion de l'urine par le rein) est inférieure à 500 ml. Elle est due soit à une réduction de la consommation de liquide, soit à un déficit de fonctionnement des reins (insuffisance rénale).

La diminution presque complète de la sécrétion des urines constitue l'**anurie**.

E.2 Caractérisation des troubles de la miction

Le nombre de fois où la personne à domicile est obligée de se lever et d'aller aux toilettes pendant la nuit permet de repérer d'éventuels troubles de la miction (nycturie, pollakiurie). Habituellement, un individu se lève de 0 à 1 fois. Les autres cas peuvent révéler la présence d'un trouble. La signification du nombre de miction d'un patient pendant une nuit est présentée sur la Figure 35.

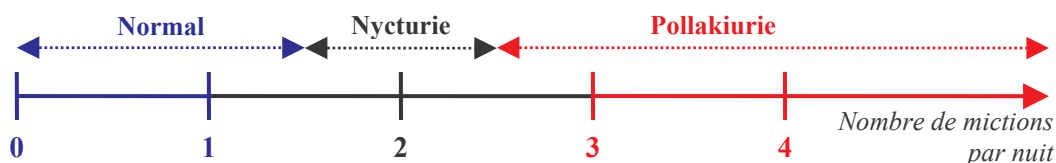


Figure 35. Echelle des valeurs du nombre de mictions par nuit

Annexe F **Glossaire médical**

<u>Adrénaline</u>	Hormone sécrétée par la médullosurrénale, par les ganglions et les fibres sympathiques post-ganglionnaires dont elle est le médiateur chimique. Sympathicomimétique, elle a une action vasoconstrictrice, broncho-dilatatrice, hypertensive : elle accélère le rythme cardiaque, ralentit le péristaltisme intestinal, élève le taux de glucose dans le sang. On l'utilise en thérapeutique surtout localement, comme hémostatique.
<u>Anémie</u>	Diminution du taux normal d'hémoglobine active dans le sang.
<u>Anthropométrie</u>	Technique de mensuration des différentes parties du corps (taille, tronc, membres, etc.) permettant, par des méthodes statistiques, de définir les caractères morphologiques d'un groupe humain.
<u>Aorte</u>	Chez les Vertébrés, artère la plus importante du corps par ses dimensions et son rôle de tronc d'origine de toutes les autres artères.
<u>Artère</u>	Vaisseau qui conduit le sang propulsé par le cœur aux tissus et aux organes.
<u>Diastole</u>	Période durant laquelle les cavités cardiaques se remplissent de sang.
<u>Glandes surrénales</u>	Les surrénales sont deux petites glandes situées au-dessus des reins. Produisant une quantité considérable d'hormones dont on distingue plus de 30 variétés, elles interviennent dans la majeure partie des fonctions organiques.
<u>Hémoglobine</u>	Pigment protéique contenu dans les globules rouges, conférant à ces derniers leurs fonctions respiratoires.
<u>Hémorragie</u>	Fuite de sang hors du système circulatoire.
<u>Hypertension</u>	Élévation anormale de la pression à l'intérieur du système circulatoire ou d'un organe comportant une partie fluide.
<u>Hypotension</u>	Diminution anormale de la pression à l'intérieur des artères ou de la cavité crânienne.
<u>Méat</u>	Canal ou orifice d'un canal. <i>Méat urétral</i> , situé à l'extrémité inférieure de l'urètre.
<u>Miction</u>	Expulsion de l'urine accumulée dans la vessie.
<u>Nycturie</u>	Émission d'urine plus abondante la nuit que le jour.
<u>Oligurie</u>	Diminution de la quantité d'urine émise en 24 heures, normalement comprise entre 800 et 1 500 ml. Habituellement, on parle d'oligurie lorsque la diurèse est inférieure à 500 ml.
<u>Pollakiurie</u>	Succession pathologique d'émissions d'urine très fréquentes et peu abondantes.
<u>Polyurie</u>	Augmentation des quantités d'urine émises par 24 heures (plus de 2 litres).
<u>Pouls</u>	Battement d'une artère superficielle correspondant à la transmission de l'onde pulsatile déclenchée par chaque contraction cardiaque.
<u>Pression artérielle</u>	La pression artérielle correspond à la résistance opposée par la paroi artérielle à la pression exercée par le flux sanguin. La pression artérielle oscille entre deux valeurs, l'une, maximale, correspondant au passage de l'ondée sanguine propagée par la contraction cardiaque systolique, l'autre, minimale, correspondant au relâchement

diastolique du cœur. Les valeurs de la pression artérielle théorique, exprimables en centimètres de mercure, peuvent être calculées, pour la maximale, en ajoutant au chiffre 10 le nombre de dizaines d'années d'âge (ex.: à 60 ans, 16), pour la minimale, en divisant la maximale par 2 et en ajoutant 1 (9 dans l'exemple précédent).

Stéthoscope

Instrument destiné à l'auscultation médicale (du thorax en particulier). Inventé par Laennec en 1819, il consistait en un simple tube acoustique placé entre l'oreille du médecin et le corps du patient. Le *stéthoscope biauriculaire*, actuellement le plus couramment employé, est constitué par une pièce réceptrice des sons, recouverte d'une membrane vibrante, à laquelle aboutissent deux tiges creuses et flexibles, terminées par des embouts se plaçant dans les oreilles de l'examineur

Systole

Première phase de la révolution cardiaque, consistant en la contraction synchrone des deux oreillettes (systole auriculaire) suivie de celle, également simultanée, des deux ventricules (systole ventriculaire).

Tissu

Association de cellules vivantes présentant des différenciations identiques ou coordonnées, de manière que leur ensemble manifeste des propriétés caractéristiques. Divers tissus sont réunis de façon définie en organes et en appareils; la coordination de leurs propriétés permet à ces organes d'assurer leurs fonctions.

Urètre

Canal excréteur évacuant l'urine du col de la vessie au méat urinaire.

Vasoconstriction

Contraction des fibres musculaires d'un vaisseau, entraînant la réduction de son calibre.

Vasodilatation

Augmentation du calibre des vaisseaux sanguins.

Annexe G Valeur diagnostique d'un signe ou d'un test [12]

Un test diagnostique est utilisé pour augmenter la certitude que l'on a sur l'identification de la pathologie présentée par un patient. On peut modéliser le processus diagnostique de la manière suivante. Avant de demander un test, le médecin associe un ensemble de probabilités *a priori* aux différentes hypothèses diagnostiques envisagées. Le résultat du test demandé permettra une révision de ces probabilités. La valeur d'un test diagnostique dépend donc de son pouvoir de séparer les différentes hypothèses en présence (pouvoir discriminant). Elle dépend aussi de son pouvoir d'induire un changement notable de thérapeutique pour le patient.

Les tests utilisés ne donnent habituellement pas une information parfaite et ne permettent pas d'identifier des pathologies avec certitude. De plus, une valeur peut différer de la vraie valeur par imprécision et/ou inexactitude. La précision fait référence à la fidélité de la mesure. On répète la mesure sur le même sujet et on obtient le même résultat. L'exactitude fait référence à la tendance des valeurs mesurées à se répartir symétriquement autour de la vraie valeur de la variable analysée.

Ces deux notions sont présentées sur la Figure 36.

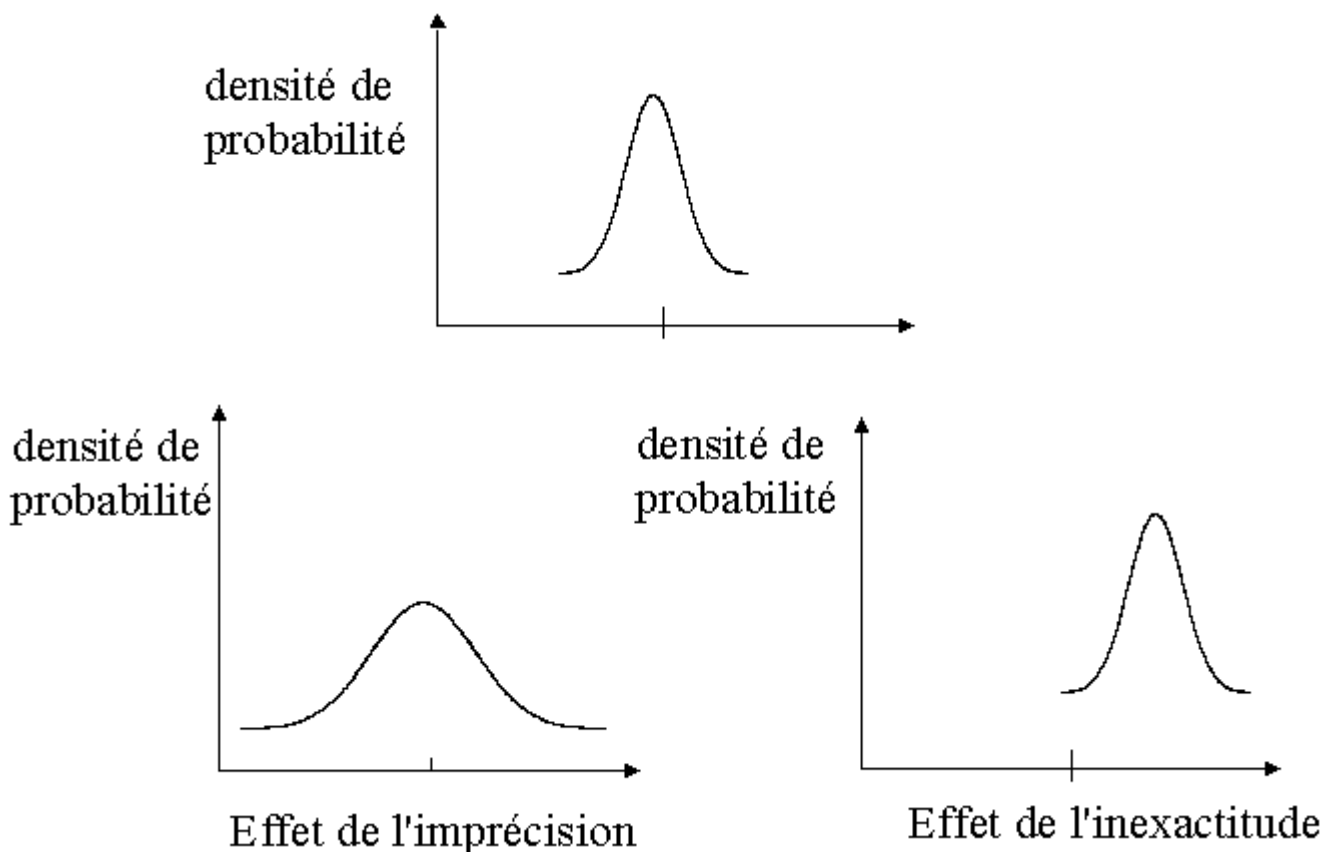


Figure 36. Représentation de l'imprécision et de l'inexactitude

La fixation des bornes de normalité de paramètres biologiques est souvent réalisée sur des critères statistiques. Au-dessus d'un certain seuil le résultat sera dit positif, en dessous, il sera négatif (Figure 37).

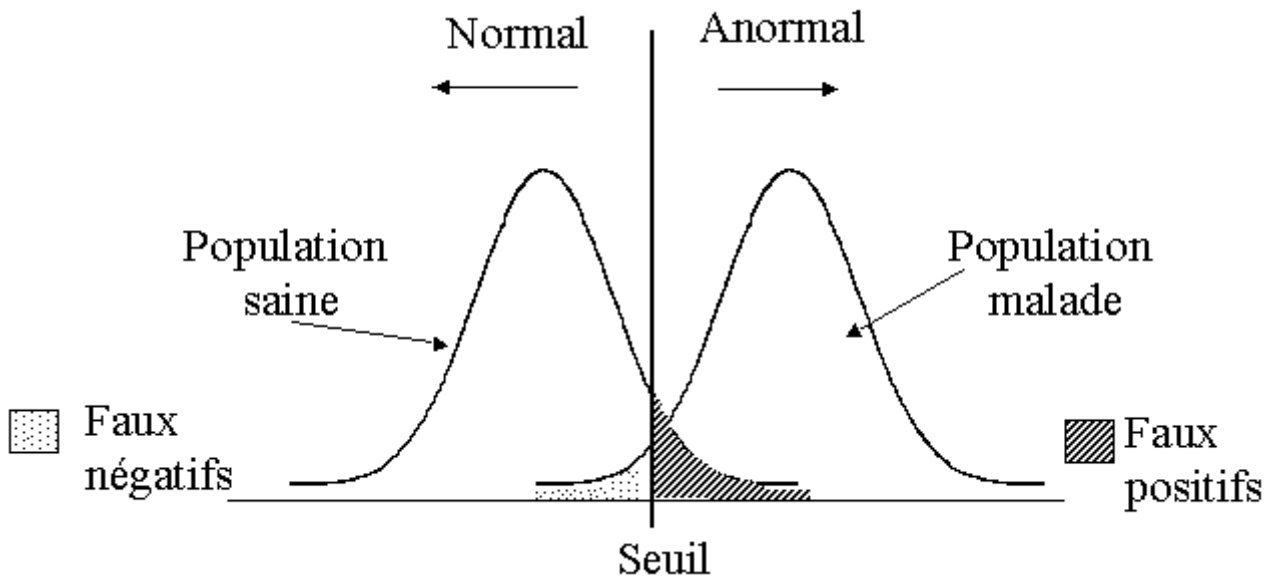


Figure 37. Borne de normalité d'un test

On peut, comme le montre le Tableau 5, identifier quatre groupes de patients :

- Les *Vrais Positifs* (VP) pour lesquels le résultat du test sera positif et la maladie présente,
- Les *Faux Positifs* (FP) pour lesquels le résultat du test sera positif et la maladie sera absente,
- Les *Faux Négatifs* (FN) qui auront un résultat du test négatif et qui présenteront la maladie,
- Les *Vrais Négatifs* (VN) qui auront un résultat du test négatif et qui ne présenteront pas la maladie.

Tableau 5. Les situations en fonction du résultat du test

	Malade	Non malade
Test positif	VP	FP
Test négatif	FN	VN

Ainsi, on peut définir le *Taux de Vrais Positifs* (ou *Sensibilité*) qui exprime la probabilité de trouver un test positif si la maladie est présente :

$$TVP = \frac{VP}{VP + FN}$$

De même, on définit le *Taux de Vrais Négatifs* (ou *Spécificité*) qui exprime la probabilité de trouver un test négatif si le patient ne présente pas une maladie :

$$TVN = \frac{VN}{VN + FP}$$

Dans le domaine médical, les mesures de la sensibilité Se et de la spécificité Sp sont fréquemment utilisées pour décrire l'exactitude d'un test clinique [6].

Il y a antagonisme entre sensibilité et spécificité qui dépendent du seuil choisi. L'amélioration de la spécificité se fait au dépend de la sensibilité et vice versa. L'exemple des tableaux 4 et 5 illustre cette remarque. Pour classer des patients dans l'une des deux catégories atteint d'une maladie M ou non atteint $\{M\}$, on utilise le test biologique du dosage de l'hémoglobine.

Sur 1000 sujets, si l'on choisi un seuil de décision à 11 g/l, on obtient les résultats du Tableau 6.

Tableau 6. Choix de M pour un seuil de décision à 11 g/l d'hémoglobine

	M	non M
Hémoglobine ≤ 11 g	100	50
Hémoglobine > 11 g	50	800

Les calculs de sensibilité Se et spécificité Sp donnent :

$$S_e = \frac{100}{150} = 0,66 \quad \text{et} \quad S_p = \frac{800}{850} = 0,94$$

La spécificité peut être considérée comme excellente, mais la sensibilité est faible.

Si l'on modifie au contraire la valeur du seuil en le fixant à 12 g/l, on obtient le Tableau 7.

Tableau 7. Choix de M pour un seuil de décision à 12 g/l d'hémoglobine

	M	non M
Hémoglobine ≤ 12 g	140	100
Hémoglobine > 12 g	10	750

Les calculs de sensibilité Se et spécificité Sp donnent :

$$S_e = \frac{140}{150} = 0,93 \quad \text{et} \quad S_p = \frac{750}{850} = 0,88$$

Ceci donne une très bonne sensibilité au détriment d'une spécificité plus faible.

Dans la pratique et pour un test donné, on peut ainsi être amené à privilégier la sensibilité ou la spécificité suivant la stratégie choisie : sélectionner les malades ou écarter les non malades. Le choix de la meilleure combinaison sensibilité-spécificité peut être aidé par le tracé de la courbe ROC (*Receiver Operating Characteristic*) [52]. Celle-ci est obtenue en représentant en ordonnée la sensibilité et en abscisse le complément à 1 de la spécificité (Figure 38). Le test qui a le meilleur pouvoir discriminant est celui qui correspond à la courbe la plus haute. Cette approche ne tient cependant pas compte des coûts respectifs des deux erreurs diagnostiques possibles : passer à côté de la maladie M ou suggérer ou poser par erreur le diagnostic M . Ces coûts ne peuvent être appréciés que sur des bases médicales, sociales ou déontologiques.

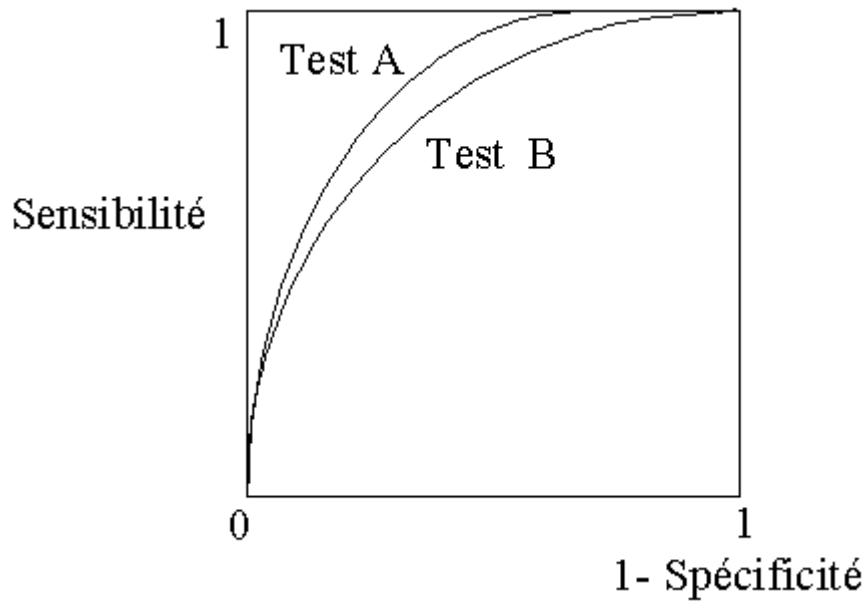


Figure 38. Courbe ROC

Ainsi, le théorème de Bayes fournit un moyen de calculer la probabilité *a posteriori* d'un diagnostic en utilisant sa probabilité *a priori* ainsi que la sensibilité et la spécificité d'un test. L'utilisation du théorème de Bayes pour interpréter une séquence de tests est licite si certaines conditions sont réalisées. On peut en effet utiliser la probabilité *a posteriori*, obtenue après le résultat du premier test, comme la probabilité *a priori* pour le second test s'il y a indépendance conditionnelle entre les deux tests. L'indépendance conditionnelle signifie que la probabilité d'obtenir un résultat avec un second test ne dépend pas du résultat du premier test utilisé. Les maladies dont on cherche à évaluer les probabilités doivent également être mutuellement exclusives

Annexe H Modèle objet de l'application Java™ développée

Le modèle de l'application Java™ développée pour l'implémentation d'un système de décision sur l'état d'une personne à domicile est constitué d'un ensemble de classes et d'interfaces. Les objets activés lors du déroulement de l'application sont des instances de ces classes.

Les paragraphes suivants décrivent l'interface utilisée pour la définition des constantes de l'application ainsi que l'ensemble des classes, leurs attributs et fonctionnalités. Enfin, les interactions entre objets lors du déroulement du processus d'apprentissage et de décision sont présentées par des diagrammes de collaborations.

H.1 Définition des constantes

Les constantes utilisées dans l'application sont définies dans l'interface *AppliConstants*. Cette interface contient les constantes utilisées pour définir ou paramétrer différentes fonctionnalités de l'application :

- les paramètres généraux de l'application, par exemple :
 - *NB_DISCRETE_CASES (3)* : nombre d'intervalles de segmentation des intervalles de variation des variables d'état à valeur dans un ensemble continu ;
 - *WIN_SIZE (20 ou 30)* : taille de la fenêtre d'observation pour le calcul des distances d'une situation au modèle d'évolution ;
 - *NULL_PROB (0,001 ou 0,00001)* : valeur de substitution utilisée pour le remplacement des probabilités nulles dans les paramètres du modèle ;
- les variables utiles à la définition de valeurs par défaut ou de valeurs d'erreur, par exemple :
 - *NO_PROB_COMPUTED (-1)* : probabilité non définie dans les paramètres du modèle ou bien dans les distances calculées au modèle. Cette valeur est utilisée par défaut pour l'initialisation des structures contenant ces paramètres ;
 - *NO_STATE_NUM (-1)* : numéro non défini pour un vecteur d'état. Ce numéro correspond à un état de la chaîne de Markov sous-jacente au modèle.
- les paramètres du codage des variables d'état du patient :
 - *NORMAL (0)* pour une valeur habituelle ;
 - *LOWER (1)* pour un valeur diminuée ;
 - *HIGHER (2)* pour une valeur augmentée ;
- des paramètres spécifiques à l'expérimentation, comme les paramètres définissant les types de chaînes de Markov expérimentés :
 - *ALL_STATES (0)* : tous les états possibles du patient sont considérés dans la chaîne de Markov (modèle 1) ;
 - *GROUP_STATES (1)* : les états sont regroupés en quatre classes selon le nombre de variables d'état ayant une valeur en dehors de l'intervalle habituel de variation (modèle 2) ;
 - *GROUP_TRANS_STATES (3)* : on regroupe les valeurs « diminuées » et « augmentées » pour ne considérer plus que deux états possibles pour chacune des variables d'état : « habituel » et « inhabituel » (modèle 3).
- les chemins d'accès aux fichiers d'entrée contenant les données de la modélisation et de l'expérimentation :
 - *INPUT_TRAINING_DATA* : chemin d'accès au fichier des données d'apprentissage ;
 - *INPUT_CHECKING_DATA, INPUT_VALIDATING_DATA* : chemin d'accès aux fichiers des données de validation des paramètres du modèle ;
 - *INPUT_TESTING_DATA1, INPUT_TESTING_DATA2, etc.* : chemin d'accès aux fichiers des données des différents tests de l'expérimentation ;

- les chemins d'accès aux fichiers de sortie utilisés pour l'écriture des résultats (probabilités d'adéquation au modèle) :
 - *OUTPUT_PROB_MI_FILE* : chemin d'accès au fichier pour l'écriture des probabilités d'adéquation au modèle 1 calculées à partir des données d'apprentissage (*INPUT_CHECKING_DATA*) ;
 - *OUTPUT_PROB_VALID_MI_FILE* : chemin d'accès au fichier pour l'écriture des probabilités d'adéquation au modèle 1 calculées à partir des données de validation (*INPUT_VALIDATING_DATA*) ;
 - *OUTPUT_PROB_TEST1_MI_FILE*, *OUTPUT_PROB_TEST2_MI_FILE* : chemin d'accès au fichier pour l'écriture des probabilités d'adéquation au modèle 1 calculées à partir des données des différents tests (*INPUT_TESTING_DATA1*, *INPUT_TESTING_DATA2*, etc.).

H.2 Description des classes

H.2.1 Classe *Telecare*

La globalité de l'application est une instance de la classe *Telecare* définissant la méthode principale *main* qui est la première invoquée. Elle contient une application, c'est-à-dire une instance *appli* de la classe *TelecareApplication*, et une interface graphique, instance de la classe *TelecareIUG*.

H.2.2 Classe *TelecareIUG*

La classe *TelecareIUG* contient l'ensemble des méthodes nécessaires à l'affichage des éléments graphiques de l'application : les différents onglets et leur contenu. Elle contient une référence de l'objet *appli*, de la classe *TelecareApplication*, et fait appel à certaines de ses méthodes pour récupérer les données à afficher.

H.2.3 Classe *TelecareApplication*

La classe *TelecareApplication* contient un attribut *myPatient* de la classe *Patient*. Cette classe gère l'ensemble des processus spécifiques au patient : l'apprentissage du modèle d'évolution et son expérimentation à partir d'un ensemble de données contenues dans des fichiers texte.

H.2.4 Classe *Patient*

La classe *Patient* représente la personne télésurveillée à domicile. Les attributs principaux de cette classe sont les suivants :

- Les particularités du patient : nom (*name*), prénom (*firstName*) et date de naissance (*birthDate*).
- Le nombre de variables d'état considérées pour ce patient (*nbStateVar*) ainsi qu'un tableau d'objets, instances de la classe *StateVariable*, définissant les caractéristiques de chaque variable d'état considérée.
- Un ensemble d'objets de la classe *StateVectors* représentant les différentes données relatives au patient prises en compte lors de l'apprentissage du modèle et de l'expérimentation. On considère notamment :
 - Des vecteurs d'apprentissage des paramètres du codage des variables d'état et du modèle d'évolution habituelle du patient (*stateVectTrain*) ;
 - Des vecteurs de validation de ces paramètres (*stateVectCheck*, *stateVectValid*) ;
 - Des vecteurs dédiés à l'expérimentation de différents types d'évolution (*stateVectTest1*, *stateVectTest2*, ...).
- Chaque objet de la classe *StateVectors* est lié à un objet de la classe *StateVectorsManager*. Sa fonction est de gérer le calcul des probabilités d'adéquation des successions de vecteurs d'état au modèle d'évolution. Il s'agit notamment de contrôler le déplacement de la fenêtre d'observation sur la suite de vecteurs d'état.

- Enfin l'objet *myPatient* contient un objet *model* de la classe **MarkovModel** représentant le modèle d'évolution habituelle du patient.

H.2.5 Classe *StateVariable*

La classe **StateVariable** gère les caractéristiques d'une variable d'état donnée. Elle comprend les attributs suivants :

- des propriétés générales : une référence de la variable (*ref*), son nom (*name*) et son unité de mesure (*unit*) ;
- des propriétés statistiques liées au patient : la moyenne (*av*) et l'écart-type (*dev*);
- les paramètres du codage : bornes inférieure (*minBound*) et supérieure (*maxBound*) de l'intervalle de segmentation « habituel » des valeurs de cette variable.

Les méthodes de cette classe permettent d'accéder à toutes ces propriétés et éventuellement de les modifier.

H.2.6 Classe *StateVectors*

La classe **StateVectors** permet donc de gérer un ensemble de données, c'est-à-dire une suite de vecteurs d'état du patient. Elle contient les attributs suivants :

- le chemin d'accès au fichier (*dataFilePath*) permettant d'aller lire les données relatives au patient ;
- le nombre de vecteurs d'état contenus dans ce fichier (*nbVect*) ;
- le nombre de variables d'état considérées (*nbStateVar*) ainsi qu'un tableau de référence de ces variables d'état (*stateVarRef*) ;
- un tableau (*vect*) à deux dimensions ou matrice (*nbStateVar* lignes et *nbVect* colonnes) contenant l'ensemble des données brutes contenues dans le fichier ;
- une matrice de même taille que la précédente (*vectDiscrete*) contenant l'ensemble des données codées en fonction des intervalles de segmentation définis pour chaque variable d'état. Ce tableau ne contient que trois valeurs possibles définies comme des constantes pour l'application dans l'interface *AppliConstants* : *NORMAL*, *HIGHER* ou *LOWER* ;
- une matrice (*vectDiscreteTranspose*) correspondant à la transposée de la matrice précédente contenant les données codées. Cette matrice transposée comprenant donc *nbVect* lignes et *nbStateVar* colonnes est utilisée par commodité dans le codage de chaque vecteur d'état en un numéro d'état de la chaîne de Markov ;
- un vecteur (*vectStates*) de dimension *nbVect* contenant le codage de chaque vecteur d'état en un numéro de la chaîne de Markov ;
- un objet *transObj* de la classe **StateTranslation** utilisé pour la conversion des vecteurs d'état en un numéro d'état de la chaîne de Markov considérée (modèle utilisant tous les états possibles ou non) ;
- un objet *stat* de la classe **Statistics** permettant le calcul des différentes propriétés statistiques des suites de vecteurs d'état :
 - *stat.average()* pour le calcul de la moyenne,
 - *stat.standardDeviation()* pour le calcul de l'écart-type ;

Cette classe fait partie de la librairie JNL (*Java Numerical Library*) [53] conçue et réalisée par *Visual Numerics*. Son utilisation requiert l'inclusion du *package VisualNumerics.math* (*import VisualNumerics.math.**).

H.2.7 Classe *StateVectorsManager*

Les objets de la classe *StateVectorsManager* gèrent le calcul des probabilités d'adéquation des successions de vecteurs d'état au modèle d'évolution, en contrôlant notamment le déplacement de la fenêtre d'observation sur la suite des vecteurs d'état. Les attributs de cette classe sont les suivants :

- un objet *stateVect* de la classe *StateVectors* référençant la succession des vecteurs d'état gérés par cette classe ;
- un paramètre *windowSize* dont la valeur est celle de la constante *WIN_SIZE* définissant la largeur de la fenêtre d'observation utilisée ;
- un index *curFirstStateIdx* dont la valeur correspond au numéro du premier état à considérer dans la succession des vecteurs d'état lors de l'incrémentement du processus de calcul de la probabilité d'adéquation d'une fenêtre d'observation au modèle ;
- un index *maxFirstStateIdx* correspondant à la valeur maximale que peut prendre la variable *curFirstStateIdx* compte-tenu de la largeur de la fenêtre d'observation considérée et du nombre de vecteurs d'état considérés.
- un vecteur *computeProb* de dimension $(maxFirstStateIdx + 1)$ permettant de stocker les probabilités d'adéquation au modèle calculées pour l'ensemble des fenêtre d'observation sur la succession des états considérée.

Les méthodes de cette classe réalisent le calcul des probabilités, permettent d'accéder aux résultats obtenus et de les écrire dans un fichier texte.

H.2.8 Classe *MarkovModel*

Cette classe représente le modèle de l'évolution habituelle du patient par une chaîne de Markov. Les attributs de cette classe correspondent aux propriétés de la chaîne de Markov et à certaines caractéristiques des données d'apprentissage :

- la variable *nbStates* représente le nombre d'états de la chaîne de Markov ;
- la variable *nbVectTrain* représente le nombre de vecteurs utilisés pour l'apprentissage des paramètres du modèle ;
- le vecteur *initProb* de dimension *nbStates* est le vecteur des probabilités initiales de chacun des états du modèle ;
- la matrice carrée *transProb* de dimension *nbStates* est la matrice des probabilités de transition entre les états.

Les méthodes de cette classe permettent d'initialiser les paramètres du modèle à partir d'une succession de vecteurs d'état dédiés à l'apprentissage, d'accéder ensuite aux propriétés de la chaîne de Markov ainsi définie et de calculer la distance d'une succession d'états au modèle.

H.2.9 Classe *StateTranslation*

Cette classe permet de gérer le lien entre les caractéristiques des vecteurs d'état (nombre de variables d'état, intervalles de segmentation de leurs valeurs) et le type de modèle considéré (tous les états possibles ou bien regroupement de certains états), c'est-à-dire la manière dont on décide de définir les états de la chaîne de Markov à partir des vecteurs d'état. Elle réalise donc le lien entre les classes *StateVectors* et *MarkovModel*. Cette classe comprend les attributs suivants :

- la variable *nbStateVar* représente le nombre de variables d'état considérées ;
- la variable *nbDiscreteCases* représente le nombre d'intervalles de segmentation des intervalles de variation des variables d'état ;

- la variable *nbGroupDiscreteCases* représente la même chose pour le cas où plusieurs classes de situations sont regroupées pour simplifier le modèle : regroupement des intervalles « augmenté » et « diminué » en une même classe d'état « inhabituel » ;
- la matrice *conv* contient les valeurs utiles au calcul du numéro de l'état (cf. 3.4.6) ;
- la matrice *convGroup* contient le même type de valeurs pour le cas où certains états ont été regroupés ;
- la variable *transType* indique le type du modèle considéré. Cette variable ne peut prendre que trois valeurs possibles définies comme des constantes pour l'application dans l'interface *AppliConstants* : *ALL_STATES* (0), *GROUP_STATES* (1) ou *GROUP_TRANS_STATES* (3).

Les méthodes de cette classe permettent de calculer selon le type de modèle le nombre d'états de la chaîne de Markov et de convertir un vecteur d'état (codé en fonction des intervalles de segmentation) en un numéro d'état de la chaîne de Markov ou l'inverse.

H.3 Interactions entre les objets

Le diagramme de la Figure 39 représente les interactions détaillées entre tous les objets de l'application impliqués dans le processus d'apprentissage et d'expérimentation des paramètres du modèle. On utilise pour cette description une représentation par des diagrammes de collaboration définis dans la notation *UML (Unified Modeling Language)* [45] développée par l'*OMG (Object management Group)*.

Les variables qui apparaissent dans le diagramme sont définies ci-dessous :

- *n* : nombre de variables d'état considérées ;
- *N* : nombre d'états de la chaîne de Markov ;
- *M* : nombre de vecteurs d'état pris en compte pour l'apprentissage ;
- *k* : nombre maximal d'incrémentes possibles lorsqu'on fait glisser la fenêtre d'observation sur la succession de vecteurs d'état de taille *M*.

Les numéros indiquant l'ordre d'envoi des messages sont hiérarchisés à la manière des titres d'un document. L'objet *:Patient* est l'émetteur de chaque série de messages et chacun de ces messages porte donc un simple numéro *x*. Si l'objet récepteur doit lui aussi faire appel aux services d'un ou plusieurs autres objets, il numérote alors la série de messages émis par le numéro du message reçu suivi d'un numéro *y*, soit *x.y*. Et ainsi de suite à chaque niveau de collaboration.

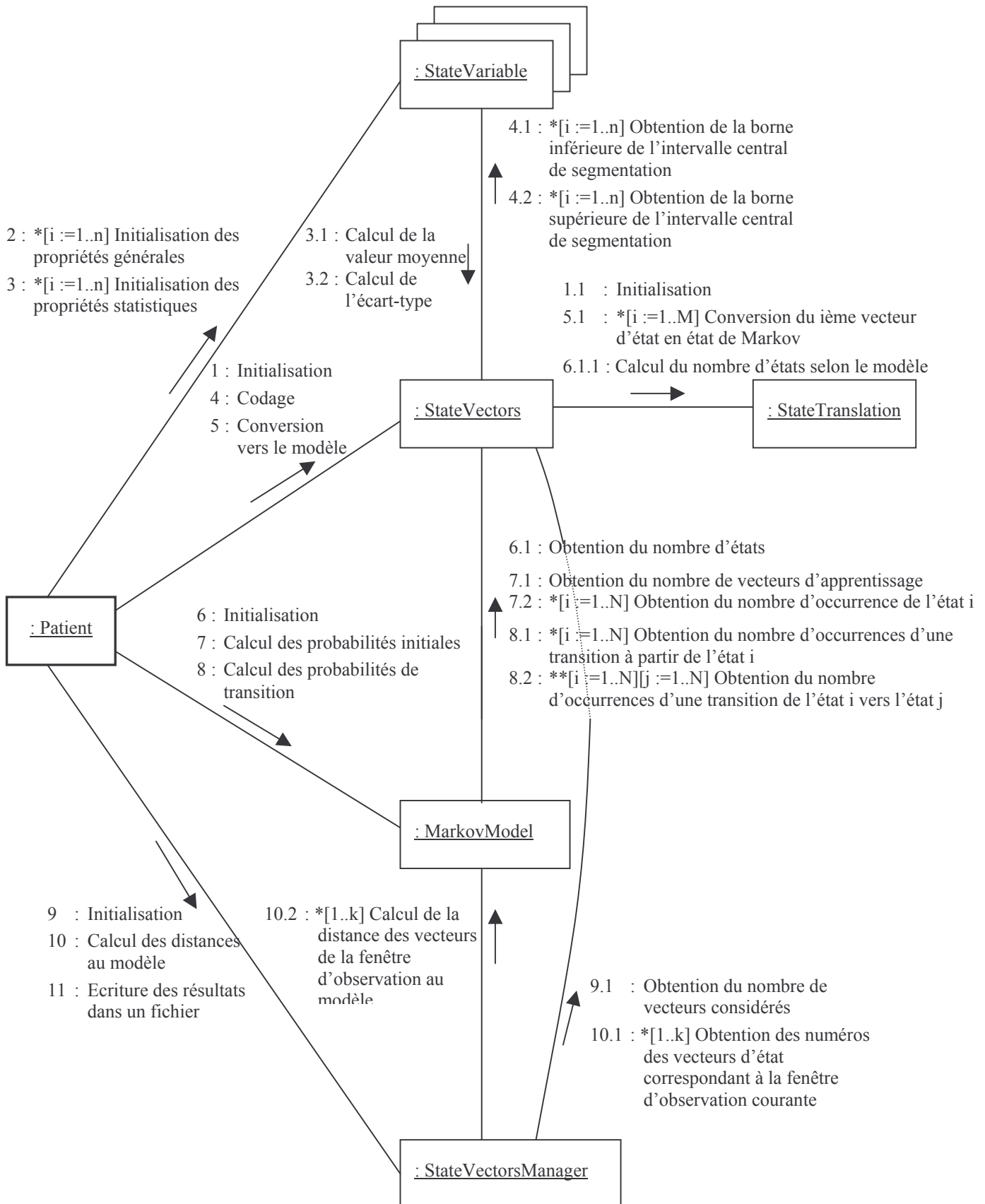


Figure 39. Représentation détaillée des interactions entre les objets de l'application

Les diagrammes de collaboration de la Figure 40 à la Figure 43 détaillent les interactions du diagramme présenté précédemment en spécifiant les arguments des messages.

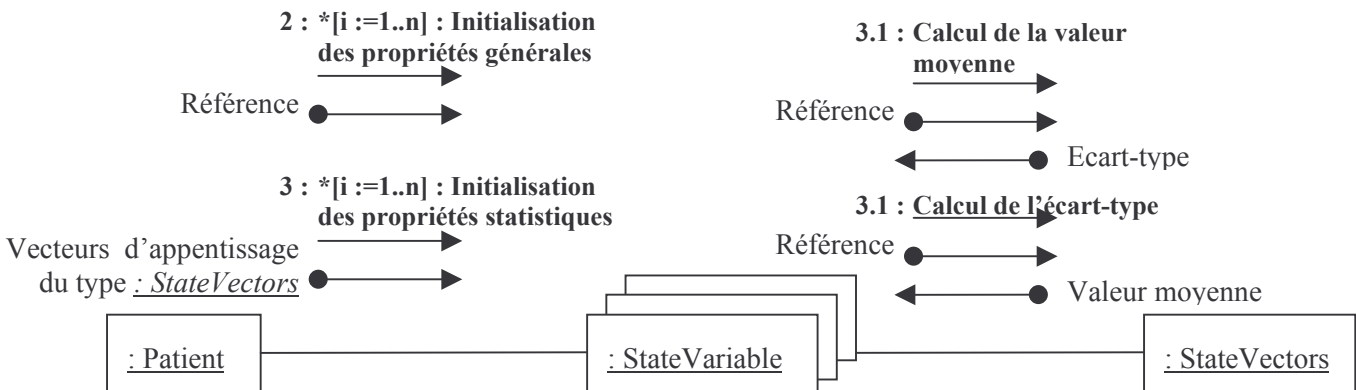


Figure 40. Diagramme de collaboration entre les objet des classes Patient et StateVariable, impliquant la collaboration d'un objet de la classe StateVectors

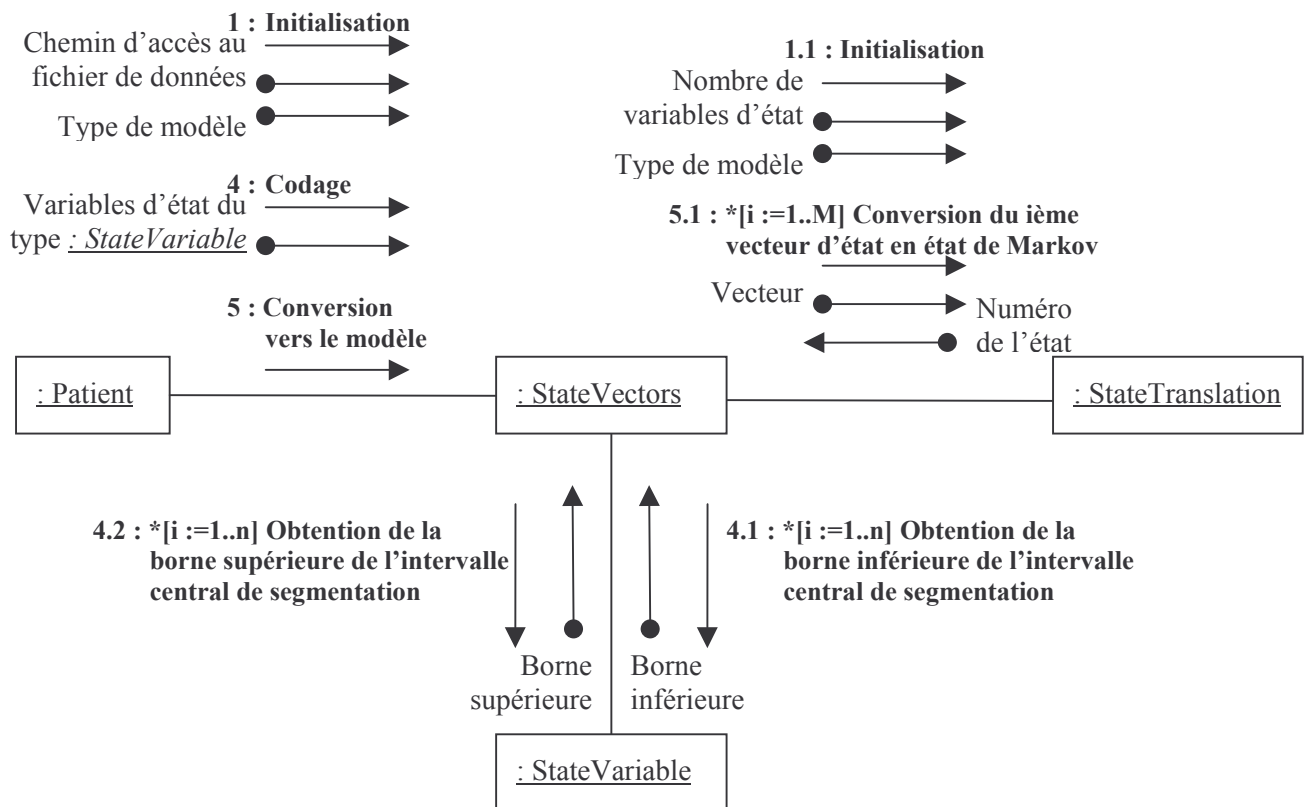


Figure 41. Diagramme de collaboration entre les objet des classes Patient et StateVectors, impliquant la collaboration d'objets des classes StateTranslation et StateVariable

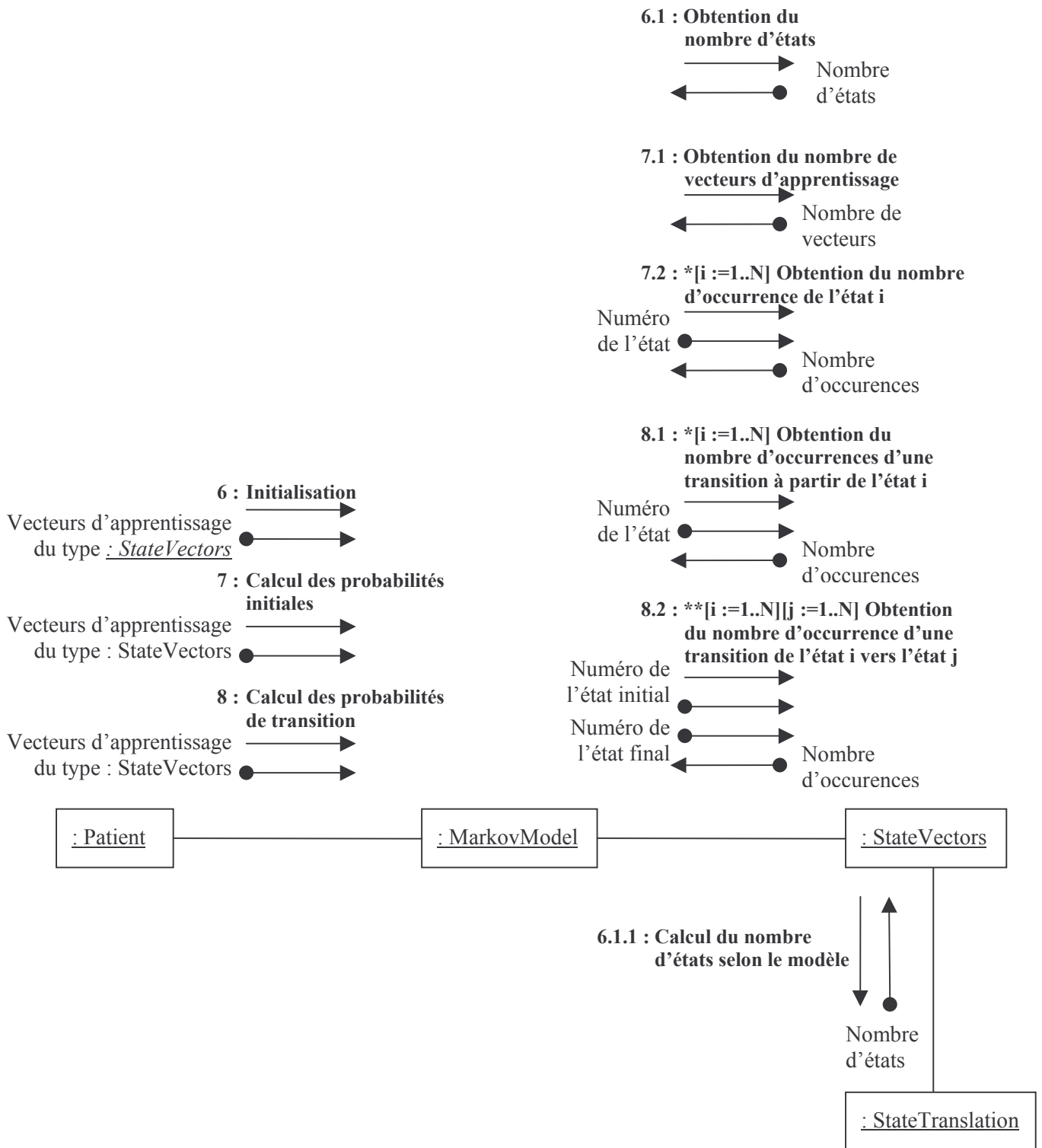


Figure 42. Diagramme de collaboration entre les objet des classes Patient et MarkovModel, impliquant la collaboration d'objets des classes StateVectors et StateTranslation

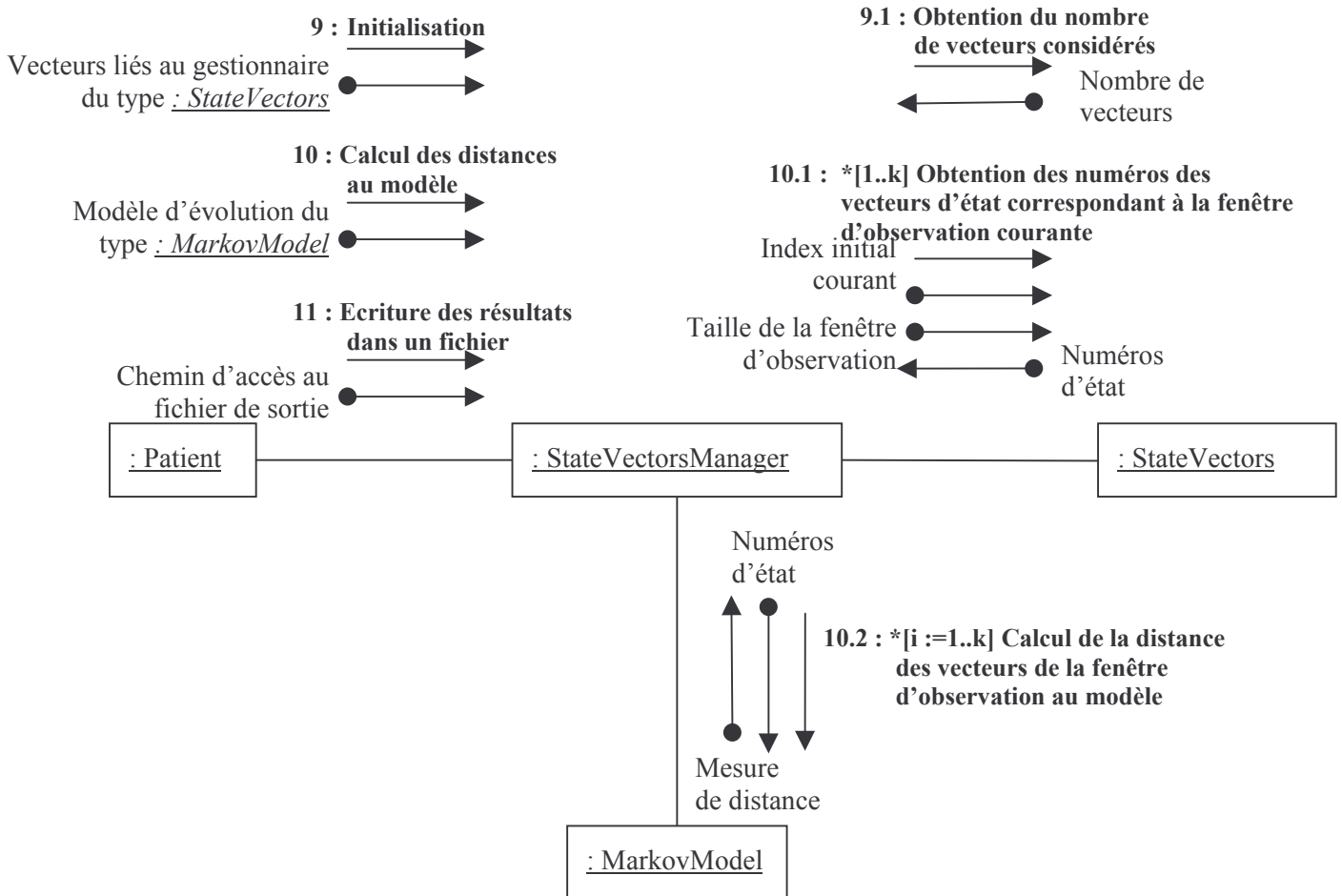


Figure 43. Diagramme de collaboration entre les objet des classes *Patient* et *StateVectorsManager*, impliquant la collaboration d'objets des classes *StateVectors* et *MarkovModel*

Annexe I Données disponibles pour l'expérimentation

Les paragraphes suivants présentent les données disponibles dans le cadre de l'expérimentation d'une application de modélisation des situations critiques d'une personne à domicile.

I.1 Données initiales

On dispose d'une succession de 92 vecteurs d'états : les 46 premiers sont dédiés à l'apprentissage et les 46 suivants à la validation du modèle. Chaque vecteur comprend les valeurs de trois variables d'état pour un même patient: les pressions artérielles systolique (PS) et diastolique (PD) ainsi que la fréquence cardiaque (FC). Ces données sont présentées dans le Tableau 8.

Tableau 8. Valeurs des pressions artérielles diastolique et systolique et de la fréquence cardiaque

Instants (apprentissage)	PS	PD	FC	Instants (validation)	PS	PD	FC
0	140	88	76	0	124	82	69
1	118	81	72	1	124	75	76
2	138	83	74	2	125	78	91
3	135	88	60	3	115	74	69
4	132	67	96	4	125	73	85
5	122	77	101	5	135	71	76
6	112	81	92	6	131	81	66
7	128	72	60	7	115	73	69
8	130	83	54	8	125	78	67
9	115	72	96	9	122	81	101
10	115	67	93	10	109	69	70
11	126	83	91	11	123	86	76
12	133	81	92	12	132	78	75
13	119	74	55	13	123	75	84
14	121	63	91	14	129	82	101
15	127	74	93	15	134	77	94
16	122	78	93	16	125	96	93
17	134	81	91	17	121	73	87
18	118	73	92	18	161	96	96
19	136	75	74	19	134	77	89
20	104	69	73	20	137	76	75
21	121	92	92	21	117	76	85
22	122	79	61	22	129	62	82
23	137	69	78	23	127	84	92
24	128	83	75	24	117	80	98
25	128	75	61	25	111	84	67
26	121	72	69	26	135	78	71
27	124	81	95	27	119	69	66
28	136	82	76	28	132	61	68
29	121	81	88	29	126	76	85
30	121	75	76	30	118	63	75
31	123	78	79	31	132	69	67
32	156	84	67	32	134	74	66
33	131	75	93	33	118	82	67
34	133	90	85	34	117	77	69
35	136	107	96	35	129	80	93
36	140	89	87	36	118	79	74
37	126	70	96	37	135	71	101
38	127	81	79	38	115	69	69
39	112	69	100	39	145	97	80
40	135	76	63	40	115	67	91
41	132	95	91	41	130	79	93
42	119	69	88	42	134	72	95
43	132	77	75	43	124	61	95
44	118	71	69	44	115	94	101
45	144	66	67	45	103	80	78

I.2 Données de test

En ajoutant une tendance à l'augmentation, à la diminution ou à la fluctuation, on obtient des données traduisant des évolutions inhabituelles de la personne à domicile. On calcule ainsi les données nécessaires pour les trois tests envisagés : vieillissement, accident aigu, succès thérapeutique. Les tableaux Tableau 9, Tableau 10 et Tableau 11 présentent les données représentatives de ces différentes tendances. Les données des colonnes de gauches sont calculées à partir des données d'apprentissage et celles des colonnes de droite à partir des données de validation. Les valeurs sont arrondies à l'entier le plus proche. Pour chaque type d'évolution sont également présentées les courbes de tendance des trois variables d'état considérées par rapport aux données représentatives de l'évolution habituelle du patient (Figure 44, Figure 45 et Figure 46).

Tableau 9. *Données caractéristiques du vieillissement*
PS augmente, PD augmente, FC augmente ses fluctuations

Instants (apprentissage)	PS	PD	FC	Instants (validation)	PS	PD	FC
0	140	88	76	0	124	82	69
1	118	81	72	1	124	75	76
2	138	83	74	2	125	78	91
3	136	89	59	3	116	75	68
4	133	68	97	4	126	74	86
5	123	78	102	5	136	72	75
6	113	82	93	6	132	82	65
7	130	74	58	7	117	75	67
8	132	85	52	8	127	80	65
9	117	74	98	9	124	83	103
10	117	69	95	10	111	71	68
11	128	85	93	11	125	88	74
12	136	84	95	12	135	81	72
13	122	77	52	13	126	78	87
14	124	66	94	14	132	85	104
15	130	77	96	15	137	80	97
16	125	81	96	16	128	99	96
17	138	85	95	17	125	77	91
18	122	77	96	18	165	100	100
19	140	79	70	19	138	81	93
20	108	73	69	20	141	80	71
21	126	97	97	21	122	81	90
22	127	84	56	22	134	67	87
23	142	74	73	23	132	89	97
24	133	88	70	24	122	85	103
25	133	80	56	25	116	89	62
26	127	78	63	26	141	84	65
27	130	87	101	27	125	75	60
28	142	88	70	28	138	67	62
29	127	87	94	29	132	82	91
30	128	82	69	30	125	70	68
31	130	85	72	31	139	76	60
32	163	91	60	32	141	81	59
33	138	82	100	33	125	89	60
34	140	97	92	34	124	84	62
35	144	115	104	35	137	88	101
36	148	97	95	36	126	87	66
37	134	78	104	37	143	79	109
38	135	89	71	38	123	77	61
39	120	77	108	39	153	105	72
40	144	85	54	40	124	76	100
41	141	104	100	41	139	88	102
42	128	78	97	42	143	81	104
43	141	86	66	43	133	70	104
44	128	81	59	44	125	104	111
45	154	76	57	45	113	90	68

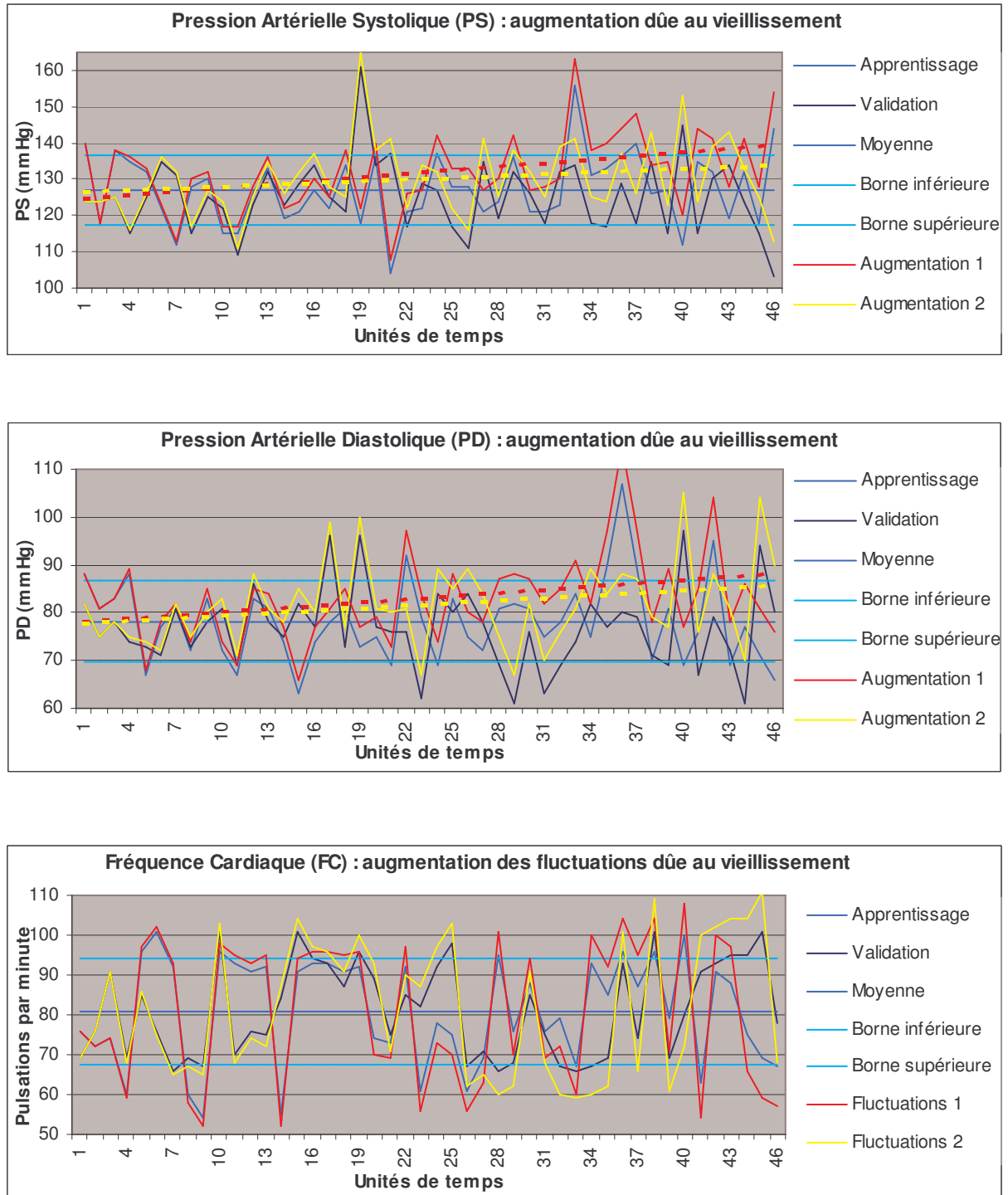


Figure 44. Courbes de tendance liées au vieillissement
 PS augmente, PD augmente, FC augmente ses fluctuations

Tableau 10. Données caractéristiques d'un accident aigu*PS augmente, PD diminue, FC reste stable*

Instants (apprentissage)	PS	PD	FC	Instants (validation)	PS	PD	FC
0	140	88	76	0	124	82	69
1	118	81	72	1	124	75	76
2	138	83	74	2	125	78	91
3	136	87	60	3	116	73	69
4	133	66	96	4	126	72	85
5	123	76	101	5	136	70	76
6	113	80	92	6	132	80	66
7	130	70	60	7	117	71	69
8	132	81	54	8	127	76	67
9	117	70	96	9	124	79	101
10	117	65	93	10	111	67	70
11	128	81	91	11	125	84	76
12	136	78	92	12	135	75	75
13	122	71	55	13	126	72	84
14	124	60	91	14	132	79	101
15	130	71	93	15	137	74	94
16	125	75	93	16	128	93	93
17	138	77	91	17	125	69	87
18	122	69	92	18	165	92	96
19	140	71	74	19	138	73	89
20	108	65	73	20	141	72	75
21	126	87	92	21	122	71	85
22	127	74	61	22	134	57	82
23	142	64	78	23	132	79	92
24	133	78	75	24	122	75	98
25	133	70	61	25	116	79	67
26	127	66	69	26	141	72	71
27	130	75	95	27	125	63	66
28	142	76	76	28	138	55	68
29	127	75	88	29	132	70	85
30	128	68	76	30	125	56	75
31	130	71	79	31	139	62	67
32	163	77	67	32	141	67	66
33	138	68	93	33	125	75	67
34	140	83	85	34	124	70	69
35	144	99	96	35	137	72	93
36	148	81	87	36	126	71	74
37	134	62	96	37	143	63	101
38	135	73	79	38	123	61	69
39	120	61	100	39	153	89	80
40	144	67	63	40	124	58	91
41	141	86	91	41	139	70	93
42	128	60	88	42	143	63	95
43	141	68	75	43	133	52	95
44	128	61	69	44	125	84	101
45	154	56	67	45	113	70	78

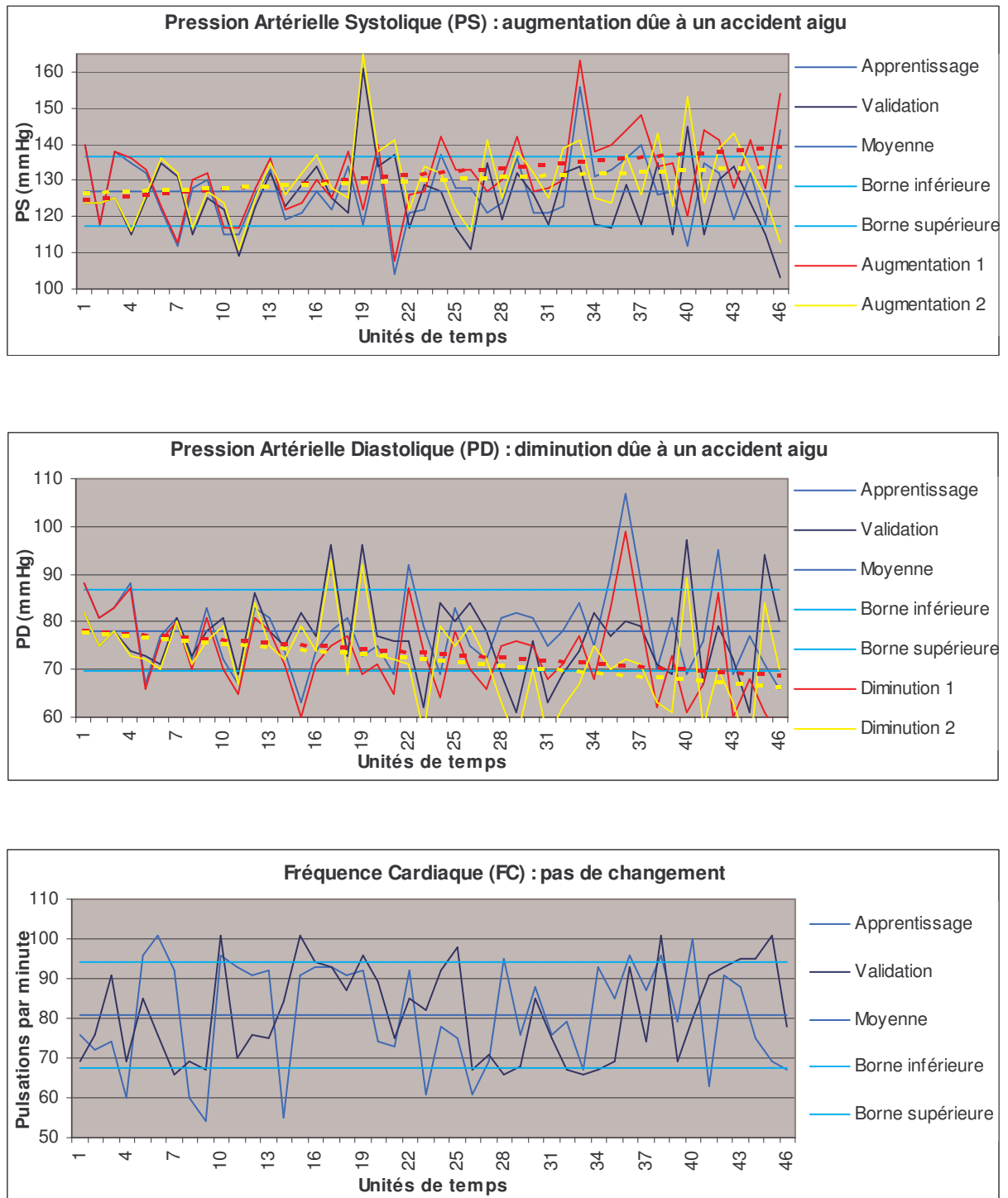


Figure 45. Courbes de tendance liées à un accident aigu
PS augmente, PD diminue, FC reste stable

Tableau 11. Données caractéristiques d'un succès thérapeutique
PS diminuée, PD diminuée, FC diminuée

Instants (apprentissage)	PS	PD	FC	Instants (validation)	PS	PD	FC
0	140	88	76	0	124	82	69
1	118	81	72	1	124	75	76
2	138	83	74	2	125	78	91
3	134	87	59	3	114	73	68
4	131	66	95	4	124	72	84
5	121	76	100	5	134	70	75
6	111	80	91	6	130	80	65
7	126	70	58	7	113	71	67
8	128	81	52	8	123	76	65
9	113	70	94	9	120	79	99
10	113	65	91	10	107	67	68
11	124	81	89	11	121	84	74
12	130	78	89	12	129	75	72
13	116	71	52	13	120	72	81
14	118	60	88	14	126	79	98
15	124	71	90	15	131	74	91
16	119	75	90	16	122	93	90
17	130	77	87	17	117	69	83
18	114	69	88	18	157	92	92
19	132	71	70	19	130	73	85
20	100	65	69	20	133	72	71
21	116	87	87	21	112	71	80
22	117	74	56	22	124	57	77
23	132	64	73	23	122	79	87
24	123	78	70	24	112	75	93
25	123	70	56	25	106	79	62
26	115	66	63	26	129	72	65
27	118	75	89	27	113	63	60
28	130	76	70	28	126	55	62
29	115	75	82	29	120	70	79
30	114	68	69	30	111	56	68
31	116	71	72	31	125	62	60
32	149	77	60	32	127	67	59
33	124	68	86	33	111	75	60
34	126	83	78	34	110	70	62
35	128	99	88	35	121	72	85
36	132	81	79	36	110	71	66
37	118	62	88	37	127	63	93
38	119	73	71	38	107	61	61
39	104	61	92	39	137	89	72
40	126	67	54	40	106	58	82
41	123	86	82	41	121	70	84
42	110	60	79	42	125	63	86
43	123	68	66	43	115	52	86
44	108	61	59	44	105	84	91
45	134	56	57	45	93	70	68

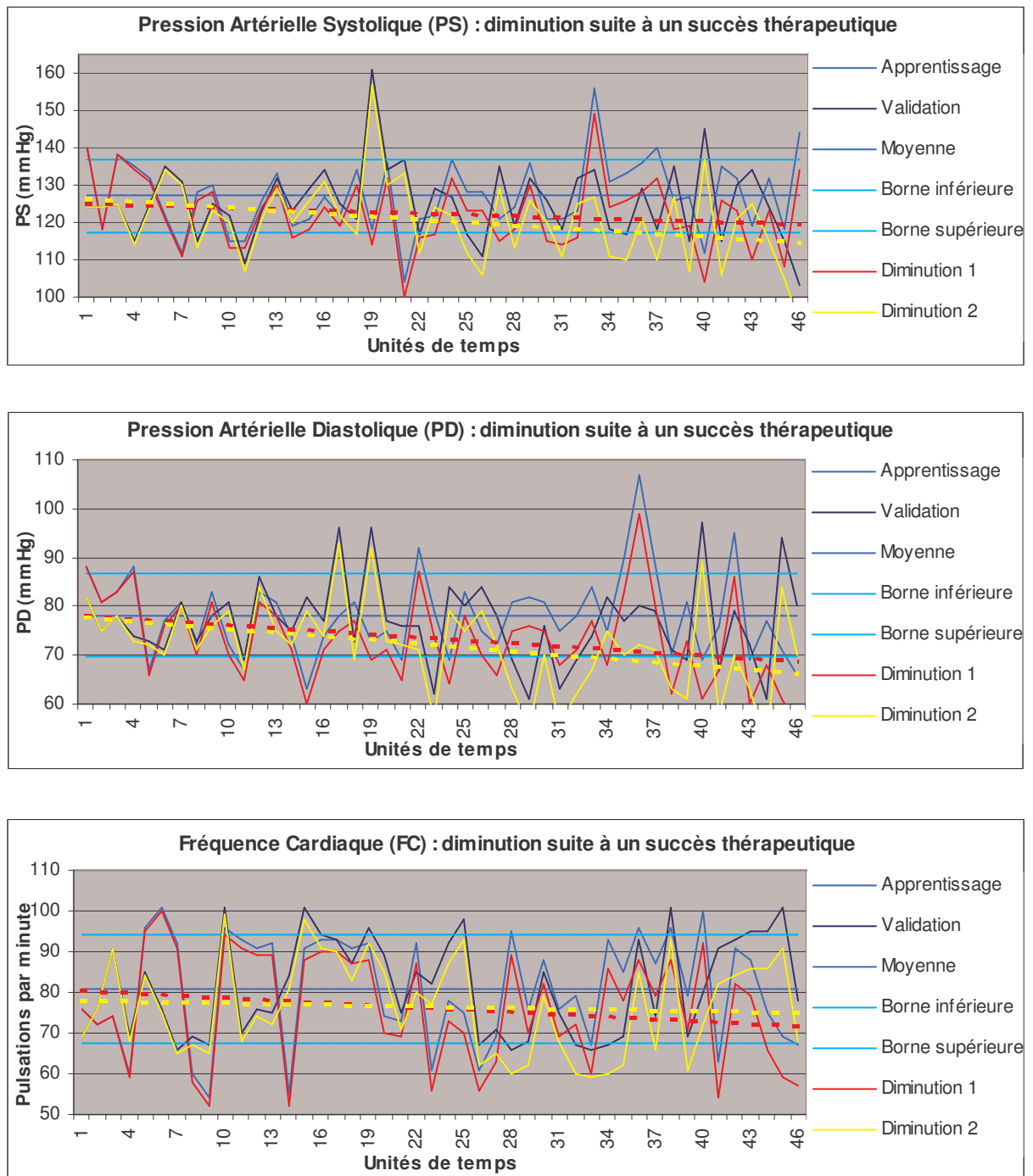


Figure 46. Courbes de tendance liées à un succès thérapeutique
PS diminue, PD diminue, FC diminue