

ENSEIGNER LE RECUEIL DE DONNÉES : ÉTUDE DE CAS EN ÉPIDÉMIOLOGIE

Marthe-Aline JUTAND¹, Karen LEFFONDRÉ², Marianne SAVÈS³
et Valérie KIEWSKY⁴

TITLE

Teaching the data collection: case studies in epidemiology

RÉSUMÉ

Bien comprendre la façon dont les données sont recueillies est un élément fondamental de la démarche statistique. Cette compréhension permet en effet une sélection adéquate de la méthode d'analyse des données et une interprétation juste des résultats. La méthode de recueil englobe non seulement l'outil et le mode de recueil, mais aussi la méthode de sélection des individus (ou unités) dans la population d'intérêt. Dans le cadre du Master mention santé publique proposé au sein de l'Université de Bordeaux à l'ISPED, l'enseignement du recueil de données, notamment dans le parcours épidémiologie, se décline selon plusieurs stratégies pédagogiques, qu'il s'agisse de cours théoriques, d'exercices d'application, d'apprentissages par projet, et de lectures critiques d'articles.

Mots-clés : enseignement de la statistique, recueil de données, questionnaire, échantillonnage, épidémiologie.

ABSTRACT

Understanding the way the data are collected is a key component of statistical thinking. A good understanding of this key component indeed allows an appropriate selection of the statistical method to be used to analyze the data, as well as an appropriate interpretation of results. Data collection methods include not only the way the data are collected from participants but also the way the subjects are selected from the population of interest. In the Master of Public Health at the University of Bordeaux, in particular in the Epidemiology stream, various teaching strategies and techniques are used to develop the data collection competences: theoretical courses, application exercises, project learning, and critical reading of articles.

Keywords: statistics education, data collection, questionnaire, sampling scheme, epidemiology.

1 Introduction

L'enseignement de la statistique en France est généralement centré sur les méthodes de traitement des données et aborde assez rarement la question du recueil de ces dernières ; ce point d'enseignement est souvent laissé à la responsabilité d'autres disciplines (économie, sociologie, psychologie, archéologie, épidémiologie, ...) que la statistique. Cependant, comme cela a été souligné en 2005 dans le rapport GAISE⁵ (Franklin *et al.*, 2005), afin de

¹ Université de Bordeaux, Laboratoire Cultures et Diffusion de Savoirs (CeDS, EA 7440), marthe-aline.jutand@u-bordeaux.fr

² Université de Bordeaux, ISPED, karen.leffondre@u-bordeaux.fr

³ Université de Bordeaux, ISPED, marianne.saves@u-bordeaux.fr

⁴ Université de Bordeaux, ISPED, valerie.kiewsky@u-bordeaux.fr

⁵ Rapport américain présentant des recommandations pour l'enseignement de la statistique.

permettre aux étudiants de développer une pensée et un raisonnement statistique, il est recommandé d'enseigner progressivement quatre étapes fondamentales de la démarche statistique que sont : 1) formuler une question, 2) collecter des données, 3) analyser les données et enfin 4) interpréter les résultats.

Il s'agit dans cet article de présenter l'approche choisie par l'équipe pédagogique du Master mention santé publique de l'Institut de Santé publique, d'Epidémiologie et de Développement (ISPED) de l'Université de Bordeaux pour permettre aux étudiants de développer les compétences nécessaires au recueil de données.

2 Place du recueil des données dans l'enseignement de la statistique

La collecte des données fait partie intégrante de l'approche hypothético-déductive assez classique dans le cadre de la démarche statistique appliquée. Sensibiliser les étudiants à l'importance d'une bonne connaissance de la construction des données, comme élément fondamental pour définir une analyse adéquate et une justesse dans l'interprétation, fait partie d'une mission de l'enseignant de statistique. L'enseignement de la statistique ne doit pas se limiter à une approche algorithmique ou calculatoire de la statistique, mais doit favoriser l'apprentissage de la démarche statistique dans son ensemble. Linda Gattuso (2011) définit la démarche statistique comme constituée des étapes suivantes : le questionnement scientifique, la planification de l'expérience (étude du contexte), la collecte des données ou la conception de systèmes d'information, l'analyse des données et leur interprétation, mais aussi la conclusion et la restitution ou la communication des résultats, et enfin le re-questionnement scientifique construit sur une approche critique et réflexive du travail et des résultats (Figure 1).

Tout ceci met en évidence l'importance d'une approche contextualisée pour l'apprentissage de la statistique appliquée et ceci tout particulièrement dans le cadre disciplinaire de l'épidémiologie. Comme le soulignaient Moore et Cobb, les données ne sont pas juste des nombres, mais des nombres dans un contexte (1997). Comprendre l'environnement de production de données est réellement nécessaire afin de pouvoir analyser correctement ces dernières. Le courant actuel des *thick-data*⁶ souligne bien le danger de travailler avec de grosses bases de données sans prendre le recul nécessaire pour apprécier l'environnement dans lequel ces données ont été recueillies. Ainsi, un fichier extrait d'une base de données collectant des données recueillies en routine, comme par exemple les bases de données de l'Insee, de l'Assurance Maladie, ou des réseaux de surveillance (de quantités particulières telles que les polluants atmosphériques ou le nombre de nouveaux cas d'une maladie donnée) doit être manipulé en en connaissant la genèse, et particulièrement la méthode du recueil des données le composant.

Nous considérons ici que les méthodes de recueil de données englobent à la fois l'outil et le mode de recueil, mais aussi la méthode de sélection des individus (ou unités) dans la

⁶ Traduit par *données épaisses* et introduisant la dimension qualitative des données et la place du contexte de la production des données, ceci en regard au courant des big data. Thématique ayant fait l'objet en 2017 du 9^e colloque Communication, Organisation, Société du Savoir et Information (COSSI) « Méthodes et stratégies de gestion de l'information par les organisations : des *big data* aux *thick data* ». <http://www.umoncton.ca/umcs-cossi/node/1>

M.-A. Jutand et al.

population d'intérêt. Car recueillir des données nécessite un environnement rassemblant les acteurs (enquêtés et enquêteurs), l'objet permettant le recueil (outil de recueil), et enfin la mise en scène permettant de mettre en relation les acteurs et l'objet (*i.e.* le mode de recueil).

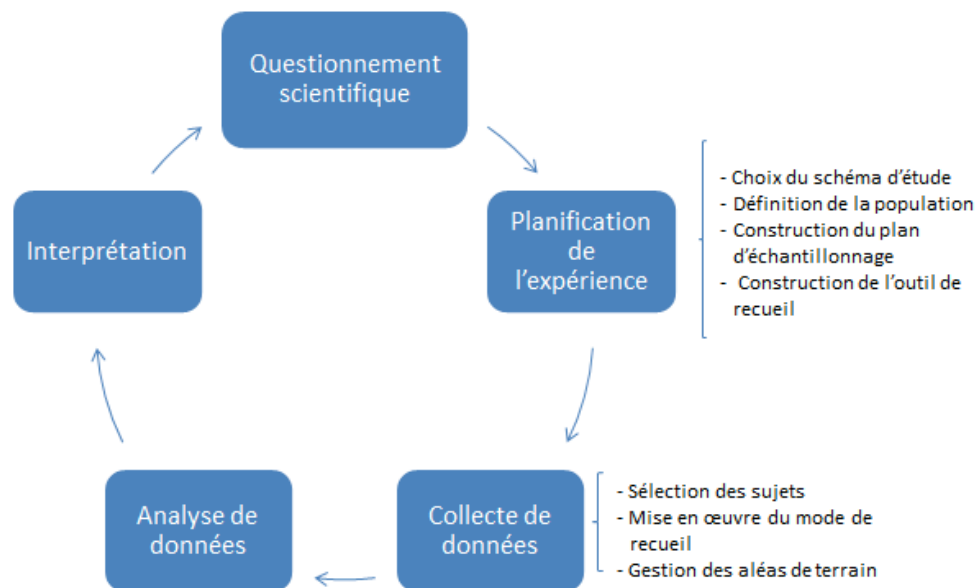


FIGURE 1 – Différentes étapes de la démarche statistique et phases importantes pour le recueil des données

Le recueil des données est de plus en plus diversifié en raison de l'évolution de nouvelles technologies (smartphone, objets connectés, ...), du développement de collectes mixtes (permettant d'associer une approche qualitative et quantitative), de la grande mobilité des sujets enquêtés (développement d'enquêtes téléphoniques, sur fixe et/ou portable, pouvant nécessiter le développement de nouvelles stratégies de sélection). Tout ceci justifie l'importance de former les étudiants à une bonne connaissance des méthodes actuellement utilisées et une capacité à comprendre de nouvelles stratégies de recueil, voire à en développer si nécessaire, au regard de l'environnement dans lequel ils seraient amenés à recueillir des informations. Il s'agit donc bien de former des étudiants ayant les compétences leur permettant de satisfaire aux situations d'élaboration d'enquêtes qu'ils rencontreront tout au long de leurs parcours professionnel.

2.1 Modes et outils de recueil et enseignement

Le mode de recueil est défini comme étant la situation dans laquelle des données sont recueillies pour une enquête, et l'outil de recueil, quant à lui, correspond au support utilisé dans ce cadre. Les différents types de mode de recueil classiques sont le face à face, l'échange téléphonique, l'envoi postal, l'envoi internet... tandis que le questionnaire, la grille d'entretien ou le cahier d'observation correspondent à des exemples d'outils de recueil.

Un même outil de recueil peut être administré selon des modes de recueil bien différents et il ne faut pas sous-estimer l'influence de ceux-ci ; par exemple, l'enquêté ne perçoit pas une question de la même façon lorsqu'il doit y répondre tranquillement chez lui ou s'il est interrogé en extérieur par un enquêteur qu'il rencontre pour la première fois.

Le mode de recueil ainsi que l'outil de recueil ont donc conjointement une influence sur l'information portée par les données. Pour s'en convaincre, il suffit de s'attarder sur les différentes manières de recueillir une information aussi simple, en apparence, que l'âge. A la lecture d'un fichier de données, le statisticien trouvera, dans la plupart des cas, une variable nommée AGE – quasiment incontournable dans les fichiers de données en épidémiologie. Cependant, cette variable peut être issue de situations bien différentes : le recueil a peut-être été réalisé par questionnaire auto-administré, la question pouvant être « quel âge avez-vous ? » ; mais il est aussi possible que la demande initiale ait été la date de naissance et qu'un calcul ait ensuite été réalisé pour indiquer l'âge dans le fichier de données ; enfin, l'information a peut-être été recueillie lors d'un face-à face entre l'enquêteur et l'enquêté, soit en posant les mêmes questions que précédemment, soit en demandant une pièce d'identité. Toutes ces situations diffèrent et peuvent conduire à des situations de réponses bien distinctes les unes des autres : il y aura plus de non-réponses dans le cas de l'utilisation de questionnaire auto-administré ou dans le cas d'absences de documents administratifs, et des réponses mensongères dans des situations de face-à face.

La connaissance du couple « mode/outil de recueil » et « fichier de données » est donc nécessaire pour réaliser correctement l'analyse statistique permettant de répondre à l'objectif de l'enquête, sans interpréter à tort certains résultats et faire un mésusage des résultats.

Dans le cadre de la stratégie pédagogique à déployer pour former les étudiants, plusieurs situations peuvent se présenter :

- soit les données ont déjà été recueillies. Il s'agira alors de sensibiliser les étudiants à l'influence de la stratégie de recueil sur les données à analyser. Pour cela, il sera intéressant de les encourager à une lecture critique du protocole d'étude, ceci afin de les sensibiliser à la nécessité d'une prise de connaissance de l'environnement de recueil de données avant de se lancer à corps perdu dans des analyses qui pourraient être *in fine* inadaptées ;
- soit les données doivent être recueillies pour répondre à l'objectif d'enquête clairement défini. Dans ce cas, il s'agira d'accompagner les étudiants dans la construction de l'ensemble des étapes du recueil, particulièrement l'outil de recueil, en précisant le comment de son usage en termes de lieux, de temps, de moyens mais aussi d'enquêtés.

2.2 Comprendre les données

Ainsi, enseigner le recueil de données peut être abordé par différentes approches qui consistent à amener les étudiants à comprendre comment le recueil a été réalisé, mais aussi à apprendre à construire la stratégie de recueil adéquate à l'objectif d'enquête. La dimension praxéologique du recueil, c'est-à-dire le passage à l'acte, est en effet fondamentale pour permettre de comprendre l'ensemble de l'information que porte chaque donnée statistique, ceci afin de ne pas limiter la donnée à un chiffre ou un nombre, mais bien de lui permettre de porter une information pertinente dans l'analyse qui en sera faite. Comme le souligne Ackoff (1989), il ne faut pas confondre données et informations ; c'est le fait de donner du sens aux données qui permet d'aboutir à l'information, et ce sens ne peut exister que s'il y a une bonne compréhension du pourquoi et du comment le recueil de ces données a été réalisé.

La lecture critique du protocole permet tout d'abord d'analyser la construction même de l'outil de recueil au regard de l'objectif de l'enquête. Accompagner l'étudiant dans une

analyse fine permettant de croiser les attendus statistiques – que nous pouvons nommer connaissances – avec les moyens mis en œuvre pour recueillir les données fait partie de cette étape. Ce travail permet de présenter à l'étudiant les différentes stratégies, mais aussi de l'aider à développer un esprit critique afin de savoir repérer les écueils à éviter.

Un second point particulièrement important à travailler avec les étudiants dans le cadre de la compréhension de l'impact du recueil des données est le choix de la population d'étude et la stratégie de sélection qui peut être réalisée.

2.3 Méthodes d'échantillonnage

Le choix du recueil de données englobe donc aussi la définition du « qui est enquêté », c'est-à-dire la définition de la population cible⁷ et du choix de la méthode d'échantillonnage des participants ou des unités statistiques. Les investigateurs doivent sélectionner le plan d'échantillonnage qui permettra de répondre au mieux à la question posée, conditionnellement aux contraintes temporelles et budgétaires de l'étude.

Il existe de nombreuses méthodes d'échantillonnage des sujets dans une population, et le choix dépend des objectifs et des contraintes de l'étude. Une étude à visée descriptive impliquera des contraintes pour l'échantillonnage différentes de celles rencontrées pour une étude à visée comparative. Les définitions claires et précises des événements étudiés et de l'objectif de l'étude sont donc les conditions *sine qua non* pour l'élaboration d'un plan d'échantillonnage adapté.

En épidémiologie, on s'intéresse souvent à la relation entre un problème de santé et une exposition à un facteur qui pourrait être une cause de ce problème de santé. Selon le type de problème de santé et l'exposition étudiés, et en fonction des contraintes temporelles et budgétaires de l'étude, on peut sélectionner les sujets avant ou après le développement du problème de santé, et on peut contrôler ou non l'exposition étudiée. On distingue ainsi les schémas d'études expérimentaux où l'exposition d'intérêt est dite « contrôlée » (c'est-à-dire imposée par l'investigateur par tirage aléatoire des sujets dans le groupe exposé ou dans le groupe non exposé), et les schémas d'études observationnels où rien n'est contrôlé si ce n'est la façon de sélectionner les sujets. Ces dernières incluent les études transversales et les études de cohorte dans lesquelles les sujets sont sélectionnés indépendamment de leur statut malade-non malade, et les études cas-témoins dans lesquelles les sujets sont sélectionnés conditionnellement à ce statut. Cette diversité de schémas d'étude induit une complexité dans la construction des échantillons, l'analyse statistique des données recueillies, et dans l'interprétation et la discussion des résultats. Il est donc nécessaire de multiplier les exemples et les applications pratiques durant le cursus de formation en épidémiologie.

3 Enseignement en épidémiologie

3.1 Epidémiologie

L'épidémiologie est l'étude de la distribution des problèmes de santé dans le temps et dans l'espace, et de leurs déterminants dans les populations humaines.

⁷ Population qui a motivé au départ la mise en place de l'enquête et que le responsable de l'enquête souhaite atteindre pour répondre à l'objectif de l'enquête.

Le cœur du métier de l'épidémiologiste est de formuler une question à propos de la fréquence ou des causes d'un problème de santé particulier, de justifier cette question à partir des données de la littérature, de proposer un schéma d'étude permettant de collecter les données nécessaires pour répondre à la question, d'organiser le recueil des données, d'analyser les données, d'interpréter les résultats et enfin de communiquer les résultats en les questionnant de manière scientifique. La démarche épidémiologique est donc étroitement liée à la démarche statistique.

Dans le cadre de l'enseignement de statistique des futurs épidémiologistes, un point est particulièrement développé : il s'agit de l'importance de l'adéquation de la méthode d'analyse statistique au schéma d'étude mis en œuvre pour recueillir les données. En effet, la méthode d'analyse des données est radicalement différente si les sujets ont été sélectionnés conditionnellement ou non à leur statut malade-non malade, par exemple. L'interprétation des résultats doit également intégrer le schéma d'étude et les biais possibles qui en découlent. Les cours de statistique appliqués à l'épidémiologie ne peuvent donc pas être déconnectés de la façon dont les données ont été recueillies.

Les formations en épidémiologie sont essentiellement dispensées au niveau du deuxième cycle universitaire en France, c'est-à-dire au niveau Master. Cette discipline attire des étudiants issus de formations initiales diverses incluant, entre autres, la médecine, la pharmacie, les sciences infirmières, la biologie, les mathématiques, la statistique, la psychologie et la géographie. Les contenus pédagogiques de Master en épidémiologie doivent donc être élaborés afin d'être accessibles à un public venant de parcours de formation très variés et n'ayant pas nécessairement de bases en statistique à l'entrée en Master 1. Pourtant, à l'issue du Master 2, tous les étudiants en épidémiologie doivent avoir acquis les bases de la démarche épidémiologique qui s'appuient entre autre sur la statistique. Les différentes étapes de cette démarche, dont le recueil de données, doivent par conséquent être enseignées progressivement tout au long des deux années de Master.

3.2 Que se passe-t-il au sein du Master mention santé publique de Bordeaux ?

A l'Université de Bordeaux, l'épidémiologie est l'un des huit parcours du Master en santé publique⁸ (1. Santé publique ; 2. Épidémiologie ; 3. Biostatistique ; 4. Santé internationale ; 5. Promotion de la santé et développement social ; 6. Santé, travail, environnement ; 7. Management des organisations médicales et médico-sociales ; 8. Système d'information et technologies informatiques pour la santé). Le premier semestre du Master 1 est commun à tous les parcours, et les cours dispensés pendant ce premier semestre doivent donc être accessibles à un public hétérogène, tant en termes de connaissances acquises que de projets professionnels dans le champ de la santé publique ; il s'agit de profils d'étudiants plus variés que les seuls postulants au parcours d'épidémiologie.

Les étudiants⁹ commencent à se spécialiser dans le parcours de leur choix au cours du second semestre du Master 1 en choisissant des unités d'enseignement (UE) optionnelles

⁸ <http://www.isped.u-bordeaux.fr/Default.aspx?tabid=110&language=fr-FR>

⁹ La première année de Master de santé publique accueille environ 100 étudiants, dont 30 en enseignement à distance, puis entre 20 et 30 étudiants sont accueillis en Master 2 d'épidémiologie, dont la moitié sont issus d'autres formations que le Master 1 de santé publique de l'ISPED-Université de Bordeaux.

appropriées. Le contenu pédagogique de chaque parcours de Master 2 est quant à lui spécifique.

Les compétences visées à l'issue du Master parcours épidémiologie sont les suivantes :

- expliquer et discuter l'apport de l'épidémiologie dans la santé publique et dans l'étude scientifique de la santé et des maladies ;
- formuler les hypothèses et les questions attenantes à une étude épidémiologique sur la base de lacunes dans la littérature scientifique ;
- proposer le meilleur schéma d'étude épidémiologique pour une question donnée, ou évaluer un schéma, en fonction du contexte, des contraintes financières, temporelles, humaines et éthiques, et des données déjà existantes ;
- sélectionner et utiliser les méthodes statistiques appropriées pour décrire et analyser les données épidémiologiques ;
- interpréter et évaluer de manière critique les résultats d'une étude épidémiologique ;
- communiquer à l'écrit et à l'oral les résultats d'une étude épidémiologique ;
- s'insérer dans un organisme de santé.

Dans le cadre de cette publication, nous nous intéressons particulièrement aux compétences centrées sur la construction des schémas d'études, l'interprétation et l'évaluation des résultats.

3.3 Les différents contenus pédagogiques

Les informations concernant le contenu du Master mention santé publique parcours épidémiologie sont accessibles à l'adresse suivante : <http://www.isped.u-bordeaux.fr/Formation/MasterenSantePublique/>.

Des éléments concernant le recueil de données sont enseignés tout au long du Master 1 ainsi que de l'année du parcours Master 2 épidémiologie, comme le précise la Figure 2. Il s'agit d'accompagner l'étudiant pour lui permettre de ne pas dissocier les données – au sens de la valeur numérique analysable d'un point de vue calculatoire – de leur contexte de recueil et donc d'interprétation.

Le processus d'apprentissage proposé durant le Master vise à permettre de développer l'autonomie de l'étudiant dans la mise en œuvre des différentes étapes de la démarche épidémiologique : exercices et projets guidés en début de formation, puis stage de fin d'étude. Parallèlement à cela, les enseignements théoriques présentent des contenus de plus en plus spécifiques et complexes, permettant ainsi une approche progressive afin d'enrichir les connaissances des étudiants pour leur donner les matériaux nécessaires à la construction de projets en épidémiologie.

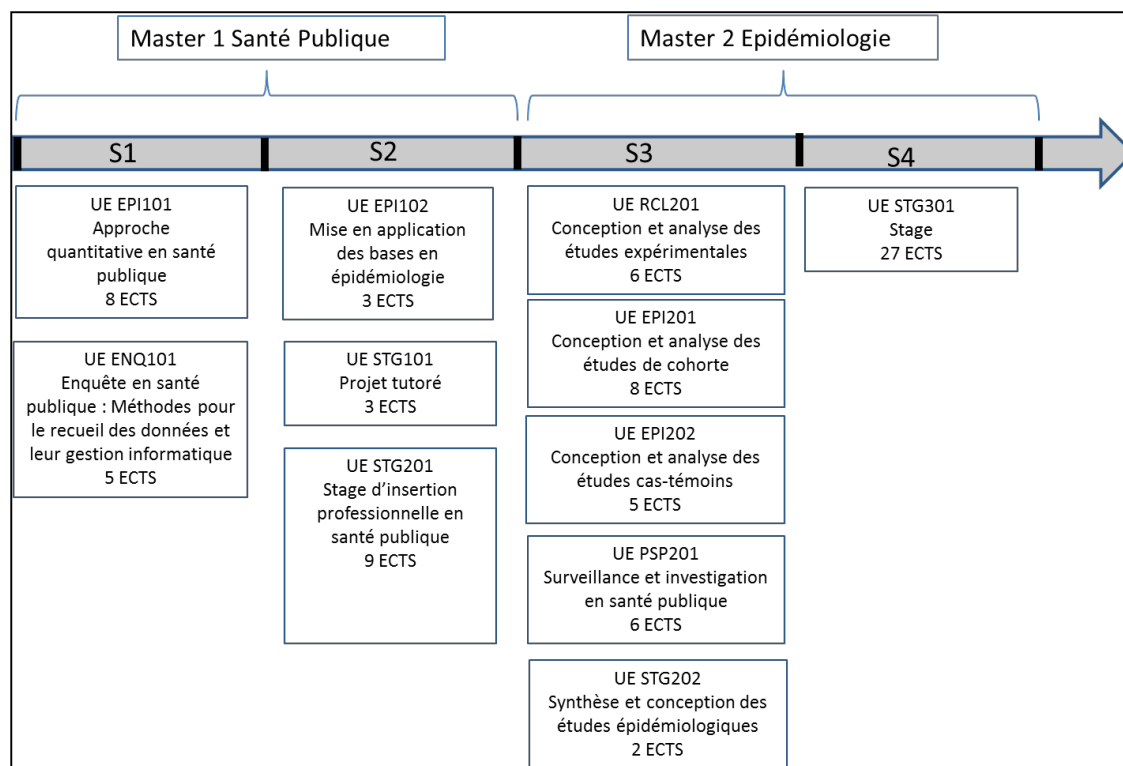


FIGURE 2 – Répartition des UE au sein desquelles le recueil de données est enseigné – Master Santé Publique Université de Bordeaux

3.3.1 Enseignement en Master 1

Dans le cadre du Master 1, les notions et contenus pédagogiques permettant d'accompagner le développement des compétences propres au recueil des données sont abordés et travaillés (cf. Figure 2) dans le cadre de :

- Deux unités d'enseignement (UE) en semestre 1 : UE EPI101 « Approche quantitative en santé publique » (8 crédits ECTS¹⁰) et UE ENQ101 « Enquête en santé publique : méthodes pour le recueil des données et leur gestion informatique » (5 crédits ECTS). L'UE EPI101 permet, entre autre, de développer les compétences quant à l'identification, la caractérisation des différents schémas d'études classiques en épidémiologie observationnelle (études transversales, études de cohorte, études cas-témoins). L'UE ENQ101 a pour objectif pédagogique de sensibiliser les étudiants au déroulement d'une enquête : du choix du sujet à la construction, la réalisation et enfin la communication des résultats.
- Deux UE au semestre 2 : UE EPI102 « Mise en application des bases en épidémiologie » et UE STG101 « Projet tutoré ». L'UE EPI102 permet d'approfondir les notions vues en semestre 1 et de comprendre, par exemple, les principes du choix des groupes comparés dans les principaux schémas d'étude. Le projet tutoré, quant à lui, a pour objectif pédagogique de permettre à l'étudiant de mettre en œuvre les connaissances acquises au cours de la formation dans le cadre d'un projet de groupe et ainsi de développer des compétences. Pour les étudiants engagés dans le parcours épidémiologie, les sujets proposés nécessitent parfois la mise en œuvre d'une enquête.

¹⁰ European credit transfert system.

- De nombreux étudiants mettent à l'épreuve leurs acquis sur le recueil de données durant le stage de fin d'année de master 1.

3.3.2 Enseignement en Master 2

Durant la seconde année du Master parcours épidémiologie, les UE sont organisées selon les grands types de schémas d'études utilisés en épidémiologie : les études expérimentales (UE RCL201), les études de cohorte (UE EPI201), les études cas-témoins (UE EPI202), et les études de surveillance épidémiologique (UE PSP201) (Figure 2). Dans chacune de ces UE, on aborde tout le déroulement spécifique des enquêtes correspondant au schéma d'étude enseigné : de la conception de l'enquête à l'analyse des données et l'interprétation/discussion des résultats. La méthode d'échantillonnage est expliquée et discutée dans le détail étant au centre même du type de schéma d'étude.

La méthode d'échantillonnage et le mode de recueil varient en effet grandement d'un type de schéma à l'autre. Par exemple, les études expérimentales (traitées dans l'UE RCL201) sont souvent menées dans le cadre d'une recherche clinique sur des populations de malades, alors que les études de cohorte (UE EPI201) sont souvent menées en population générale avec des sujets suivis sur plusieurs années et un recueil de données répété dans le temps. Dans ces études de cohorte, les questionnaires peuvent évoluer au cours du temps, et les sujets peuvent être amenés à faire des examens cliniques ou biologiques à différents moments du suivi. L'échantillon de sujets évolue lui-même dans le temps, avec des perdus de vue et des décédés, ce qui complexifie l'analyse des données et l'interprétation des résultats. Dans les études cas-témoins (UE EPI202), on échantillonne des malades (les cas) et des non malades (les témoins) au moment de l'enquête, et on les interroge sur leurs possibles expositions à des facteurs potentiellement causes de la maladie. La qualité du questionnaire et du mode de recueil (idéalement une entrevue en face-à-face par un enquêteur formé) est donc particulièrement importante dans ces études. L'échantillonnage est également très complexe dans ce type d'étude car l'échantillon des témoins doit être représentatif de la population source des cas. Comme il est généralement difficile de s'en assurer, les résultats des analyses statistiques réalisées sur les données finalement collectées doivent être discutés par rapport aux biais potentiels engendrés par l'échantillonnage réalisé. Comme pour tous les types d'étude épidémiologique, les biais potentiels engendrés par le mode de recueil des données doivent également être discutés.

Les méthodes pédagogiques déployées au sein des différentes UE associent des enseignements théoriques à des exercices dirigés, ou des apprentissages par projets dans le cadre de projets collectifs. Les étudiants sont accompagnés par des enseignants tout au long de leur formation et, selon le type de situations pédagogiques, cela se traduit par des échanges en présentiel individuellement ou collectivement, mais aussi à distance *via* des forums ou directement par échange de courriels.

4 Différentes stratégies pédagogiques

4.1 Apprendre à construire une enquête

Pour apprendre le B.A.BA de la construction d'une enquête, dès le premier semestre du Master 1, les étudiants doivent concevoir de manière progressive un protocole d'enquête, ceci pour qu'ils prennent conscience de l'importance de s'attarder sur chaque étape pour élaborer

une enquête de qualité¹¹. Puis, au cours de leur formation, ils peuvent être amenés à élaborer des protocoles et à mettre en œuvre les enquêtes dans le cadre d'une approche guidée par un enseignant ou en autonomie (avec cependant un référent) durant le projet tutoré ou les stages.

Dans cet article, les phases de choix, de justification et de rédaction de l'objectif de l'étude ne sont pas présentées, car c'est un point pouvant faire l'objet d'un article à part entière tant il s'agit d'une étape difficile pour les étudiants qui ne perçoivent pas naturellement l'importance de spécifier les attendus non pas en termes de résultats mais bien d'objectifs.

La construction d'une enquête, notamment épidémiologique, se décline réellement en trois temps (Jutand, Fagot Campagna et Ferley, 2012) :

- le temps d'élaboration : période de rédaction du protocole, nécessitant rigueur, argumentation et connaissance de la littérature scientifique sur le sujet, et connaissance des contraintes de l'environnement sociologique, scientifique, politique et économique dans lequel doit être réalisée l'enquête ;
- le temps de réalisation : période durant laquelle les enquêtés seront contactés, et les données recueillies ;
- le temps de production : période d'analyse des données, de production de résultats, de rédaction d'articles et autres communications.

Dans le cadre de l'UE ENQ101 « Enquête en santé publique » du Master 1 (Figure 2), il s'agit de visiter ces trois temps et de comprendre l'importance de la cohérence d'ensemble. A l'issue de cette UE, les étudiants doivent prendre conscience de l'importance de chaque phase pour réaliser correctement des enquêtes ; il est important de rappeler que l'enquête est un outil primordial dans le champ de la santé publique, tant pour l'étape diagnostique qu'évaluative de programme de santé publique. Il est donc important que les étudiants, débutant un parcours de santé publique, comprennent, *via* cette activité pédagogique, la pluralité des compétences nécessaires pour mener à bien une enquête et l'importance du travail collaboratif et pluridisciplinaire.¹²

Cette UE ENQ101 offre l'opportunité aux étudiants de mettre à l'épreuve le protocole qu'ils produisent, et ainsi de confronter le projet à sa réalisation. En interrogeant eux-mêmes les « vrais » participants de l'étude, les étudiants peuvent alors apprécier l'importance de a) la formulation des questions du questionnaire (fermées/ouvertes), b) l'adaptation du questionnaire à la population interrogée (jeune/âgée, éduquée/peu éduquée, active/inactive), et c) la façon dont le questionnaire est administré (auto-administration/téléphone/face-à-face).

L'UE ENQ101 permet à l'enseignant d'aborder toutes les phases d'une enquête en proposant la mise en pratique. Il s'agit d'une approche pédagogique guidée par l'enseignant qui construit de manière rigoureuse les différents temps d'apprentissage des phases de l'enquête. Plusieurs méthodes pédagogiques peuvent être utilisées par l'enseignant pour atteindre son objectif pédagogique : il peut s'agir d'un apprentissage par projet, ou d'un enseignement déroulant des phases de cours théoriques suivies par des mises en application

¹¹ C'est-à-dire une enquête permettant de recueillir des données et donc des informations permettant de construire de nouvelles connaissances satisfaisant à l'objectif d'enquête fixé initialement.

¹² Les étudiants rejoignant le Master de santé publique sont issus de parcours très variés (psychologie, droit, statistique, biologie, sociologie, formations para-médicales, ...) et il est important de leur montrer la richesse du groupe et l'intérêt de l'apport des compétences de chacun pour le développement de projets collectifs.

sur les projets, couplées aussi avec quelques travaux dirigés non directement liés au sujet de l'enquête que les étudiants construisent. L'organisation et les contraintes pédagogiques conduisent l'enseignant à définir chaque année les méthodes pédagogiques les plus adéquates et réalisables.

Le réinvestissement de cet apprentissage est possible pour certains étudiants dans le cadre du projet tutoré (UE STG101) du Master 1, avec un travail collectif qui s'organise sur une période de deux mois de fin janvier à fin mars, en parallèle aux autres cours. Les étudiants s'organisent en groupe de quatre ou cinq étudiants, et doivent répondre à une demande d'une institution de santé publique ou d'un professionnel chercheur ou non. Ils doivent donc organiser leur travail collectif et peuvent solliciter le porteur du projet pour leur permettre d'avancer. C'est donc en quasi autonomie qu'ils mettent en œuvre les acquis de la formation (voire de leurs formations antérieures). La construction d'enquêtes est une méthode fréquemment utilisée dans ce cadre. Voici quelques exemples de sujets de projets tutorés :

- décrire les pratiques professionnelles de prise en charge du tabagisme gravidique au cours des consultations prénatales, auprès des sages-femmes de la maternité de Pellegrin ;
- décrire le sommeil des externes en médecine en France durant l'année universitaire 2016-2017 ;
- décrire les connaissances, attitudes et pratiques vis-à-vis de la contraception chez les étudiantes inscrites en Licence à Bordeaux.

En Master 2, dans chaque UE RCL201, EPI201, EPI202, et PSP201, les concepts fondamentaux de l'échantillonnage des sujets et du mode de recueil propre au type de schéma d'étude sont d'abord abordés par des enseignements théoriques. Dans chacune de ces UE, il s'agit également de rendre compte de la réalité de terrain de la mise en place de telles études, en présentant des retours d'expériences par des chercheurs ou des acteurs de santé publique. Ces interventions sont généralement très appréciées des étudiants.

Enfin l'ensemble de ces compétences sont mises à l'épreuve au cours des stages que les étudiants sont amenés à réaliser à la fin du Master 1 (durée minimale de 8 semaines) et du Master 2 (durée minimale de 20 semaines).

4.2 Zoom sur la construction de l'outil de recueil

La construction de l'outil de recueil fait l'objet d'une attention toute particulière dans le cadre de l'UE ENQ101 « Enquête en santé publique », premier temps d'apprentissage concernant le recueil de données, car il s'agit d'une étape cruciale dans la réalisation d'enquêtes qui ne concerne pas uniquement les étudiants se dirigeant vers le parcours épidémiologie. Les méthodes d'enquêtes sont tout à fait transposables à d'autres domaines d'études que l'épidémiologie, par exemple la sociologie, le management ou la santé communautaire.

Construire un outil de recueil adéquat n'est pas chose facile contrairement à ce qu'il peut en paraître en première approche. Les étudiants ne prennent la mesure de la difficulté que s'ils élaborent eux-mêmes un questionnaire ou une grille d'entretien, s'ils le/la testent et s'ils recueillent, avec cet outil de recueil, les données dont ils ont besoin pour leur enquête.

Pour l'UE ENQ101, les étudiants travaillent en petits groupes de quatre ou cinq étudiants, et doivent proposer une méthode de recueil de données (mode et outils de recueil) pour

répondre à une question particulière. Les étudiants sont ainsi amenés à rédiger un questionnaire et à le tester auprès des étudiants de la même promotion travaillant sur un autre protocole d'enquête. Dans cette situation, chaque étudiant est créateur d'un outil de recueil, puis testeur. La phase de test est très formatrice pour les étudiants, tant comme testeur que comme rédacteur de l'outil de recueil. En tant que testeurs, ils doivent apprendre à réaliser une lecture critique et argumentée de l'outil de recueil, et il leur est demandé de faire des propositions d'amélioration¹³. En tant que rédacteurs, ils doivent apprendre à entendre et accepter les remarques et suggestions qui leur sont faites, et savoir prendre en considération les éléments permettant d'améliorer leur outil de recueil.

Le questionnaire proposé est donc évalué par les pairs, dans un premier temps en termes de pertinence de sa construction au regard de l'objectif, mais aussi quant à la qualité syntaxique et au choix du vocabulaire employé au regard de la population visée. Le retour de cette phase de tests réalisé intra-formation permet aux étudiants de prendre conscience de la difficulté qu'il y a à rédiger des questions claires et univoques¹⁴. Le questionnaire est également relu par les enseignants qui accompagnent les étudiants tout au long du projet avant qu'ils réalisent des tests en situations réelles, c'est-à-dire respectant le mode de recueil et ceci auprès de sujets ressemblant le plus possible au profil recherché¹⁵. Après chaque phase de test ou de relecture (par les pairs, par l'enseignant et par les sujets potentiels), les étudiants se retrouvent en petits groupes pour apporter les modifications qui leur semblent adéquates.

Cet apprentissage des différentes phases de construction d'un outil de recueil, mettant en jeu différents acteurs (les rédacteurs, les relecteurs, les testeurs) est un élément important dans la formation à la construction d'enquête tant les biais (biais d'information¹⁶) induits par un mauvais outil de recueil peuvent être importants lors de l'analyse des résultats.

4.3 Qui enquêter ?

Le choix des personnes à enquêter nécessite tout d'abord le choix du schéma d'étude (enquête de cohorte, enquête cas-témoins, ...) dépendant directement de l'objectif de l'enquête, de la définition de la population cible, du choix du mode et de l'outil de recueil et de la connaissance des contraintes associées. Tous ces éléments permettent dans un premier temps de caractériser la population source¹⁷, et de choisir la stratégie d'échantillonnage la plus appropriée à l'objectif de l'enquête.

¹³ Il s'agit de la demande pédagogique émise par l'enseignant, ceci afin que les étudiants apprennent à ne pas se limiter à des commentaires de types « c'est bien » ou « c'est incompréhensible », mais à préciser ce qui n'est pas satisfaisant et l'amélioration à y apporter.

¹⁴ Un des retours surprenant le plus les étudiants est de s'apercevoir que leurs camarades n'ont pas tous donné le même sens à la question qui leur était posée. Cette hétérogénéité de compréhension est toujours un élément entraînant de longs échanges entre les étudiants qui ont du mal à concevoir que quelqu'un n'ait pas compris ce qu'ils demandaient.

¹⁵ Les étudiants sont sensibilisés à l'importance de ne pas inclure dans l'enquête des sujets qui auraient participé au test de l'outil de recueil, pour ne pas biaiser les réponses de l'enquête.

¹⁶ Biais survenant lors du recueil des informations. Dans une enquête épidémiologique descriptive par exemple, il y a biais d'information lorsque le classement des sujets de l'échantillon vis-à-vis du statut du phénomène de santé étudié est erroné, et que cela conduit à une estimation biaisée du paramètre d'intérêt (par exemple fréquence de la maladie dans la population cible ou source).

¹⁷ Population accessible lors de l'enquête selon la stratégie définie.

M.-A. Jutand et al.

La méthode de sélection des sujets dans la population est enseignée progressivement dans différentes UE réparties sur les deux années de Master. Cette étape fondamentale de la démarche épidémiologique est enseignée au travers :

1. de présentations des différents schémas d'études en épidémiologie (au cours d'enseignements théoriques d'épidémiologie en Master 1 et Master 2) ;
2. de temps d'exercices où les étudiants sont amenés, à partir de questions précises, à critiquer des schémas d'études utilisés dans des articles publiés ;
3. de temps d'exercices durant lesquels les étudiants doivent proposer des plans d'échantillonnage au regard d'une situation proposée. Exemple d'exercice : « *On souhaite estimer les effets secondaires d'un médicament M vendu sans ordonnance en pharmacie. Pour cela une enquête doit être réalisée auprès de 1000 personnes de plus de 18 ans. Quelle méthode de sélection des sujets proposez-vous ?* » ;
4. de temps d'exercices durant lesquels les étudiants doivent réaliser eux-mêmes l'échantillonnage des sujets dans une base de données existante ;
5. de projets où les étudiants sont amenés à proposer eux-mêmes le meilleur schéma d'étude en fonction de la question posée.
 - a. En Master 1, ce type de travail est réalisé dans le cadre de l'UE ENQ101 et de l'UE projet tutoré du Master 1 ;
 - b. En Master 2, dans le cadre de projets étalés sur un mois, comme par exemple celui de l'UE EPI202 sur les études cas-témoins, les étudiants sont amenés à choisir une méthode d'échantillonnage de sujets au sein d'une base de données réelles, afin de répondre à une question de recherche bien définie, tout en respectant certaines contraintes imposées reflétant les conditions réelles (par exemple limitation de la taille d'échantillon à un nombre maximum de sujets afin de respecter les contraintes financières). Ils doivent réaliser eux-mêmes cet échantillonnage, analyser les données, et présenter et discuter les résultats à l'écrit et à l'oral devant l'ensemble de la promotion ainsi que trois à quatre enseignants. Cette restitution orale permet de revenir sur les fondamentaux, et contribue à développer la capacité des étudiants à discuter leur choix notamment sur la méthode d'échantillonnage. Cette restitution collective contribue également à développer l'esprit critique des étudiants. Dans cette UE comme dans tous les exercices dirigés et les projets des autres UE, nous nous efforçons de confronter les étudiants à des données épidémiologiques réelles, donc incomplètes et parfois erronées. Cela permet non seulement de leur enseigner des méthodes d'analyse permettant de faire face à ce type de problème, mais également de les sensibiliser à l'importance de la qualité du recueil des données ;
6. de stages de Master 1 et de Master 2 où les étudiants sont amenés à proposer eux-mêmes un schéma d'étude, ou à défendre et discuter avec un esprit critique le schéma d'étude qui leur a été imposé pour répondre à la question de recherche.

4.4 Développer la capacité de lecture critique

Au-delà de la capacité à construire des projets d'enquête, il est aussi important de développer chez les étudiants les capacités d'analyse critique de protocole et de schémas d'étude et donc d'évaluation de la qualité et de la pertinence de la méthode de recueil au regard de l'objectif de l'enquête.

En Master 2, cet esprit critique est travaillé tout au long de l'année. Ainsi, les exercices dirigés concernant le recueil des données peuvent consister par exemple à une lecture critique d'un article présentant les résultats d'une étude épidémiologique. Le recueil des données (méthode d'échantillonnage et mode de recueil) est systématiquement décrit dans ce type d'article, de manière plus ou moins détaillée, et en principe les biais potentiels engendrés par la méthode de recueil utilisée sont discutés dans la partie discussion. A partir de questions précises sur l'article, les étudiants sont amenés à discuter de ces aspects, ce qui leur permet de mettre en application les concepts abordés en cours, et développer leur esprit critique essentiel en épidémiologie.

Dans le cadre des UE STG202 et STG301, les étudiants développent leur sujet de stage (UE STG202) et effectuent leur stage de 20 semaines (UE STG301) dans une équipe réalisant des études épidémiologiques, avec un encadrant épidémiologiste. Il est accordé beaucoup d'importance au stage et de nombreuses heures enseignant sont dédiées à l'accompagnement de ces stages. Le recueil des données dans une étude épidémiologique s'étendant souvent sur plusieurs mois voire années, les étudiants participent assez rarement au recueil des données dans le cadre de leur stage. Ils arrivent en général quand le recueil est terminé, et donc à la phase d'analyse des données. Cependant, comme l'analyse des données, l'interprétation et la discussion des résultats se fait au regard de la méthode de recueil, il est exigé qu'ils maîtrisent parfaitement la façon dont les données ont été recueillies. En amont de la lecture de leur rapport de stage et de leur soutenance de stage en juin, quatre ateliers de stage qui se déroulent en octobre, décembre, janvier, et avril permettent de tester cette maîtrise en les interrogeant sur le recueil de données. Lors de ces ateliers, les étudiants présentent devant la promotion entière et quatre enseignants l'état d'avancement de leur sujet de stage. Ils doivent expliquer clairement le mode d'échantillonnage des sujets et le recueil des données de l'étude épidémiologique dont ils ont à analyser les données. Ils doivent progressivement identifier et discuter les limites de l'approche choisie par les investigateurs de l'étude, afin de pouvoir discuter l'impact potentiel de ces limites sur leurs futurs résultats. Ces limites du recueil de données sont à nouveau largement discutées dans le rapport de stage et lors de la soutenance.

Ces différents exemples illustrent les efforts de l'équipe pédagogique du Master pour développer l'aptitude à la lecture critique d'articles et/ou de protocoles. Il s'agit ainsi d'apporter et de consolider des savoirs, mais aussi de renforcer l'esprit critique, notamment vis-à-vis du recueil des données. Cet ensemble permet également aux étudiants de développer leur capacité d'expertise de documents scientifiques et de protocoles, qui est une compétence très recherchée dans le champ de la santé publique.

5 Conclusion

Enseigner la démarche de recueil des données en Master de santé publique est indispensable, et particulièrement dans le parcours épidémiologie, car, en tant que « *raisonnement et méthode (...) appliqués à la description des phénomènes de santé, à l'explication de leur étiologie et à la recherche des méthodes d'intervention les plus efficaces* » (Jenicek et Cléroux, 1982), l'épidémiologie est une science prenant appui dans de nombreuses situations sur des données recueillies par différentes techniques. Connaître les différentes méthodes permettant le recueil des données (qu'il s'agisse des outils de recueil, des modes de recueil et de leurs contraintes associées et des stratégies de construction de plan d'échantillonnage), analyser la pertinence de la mise en place de tel recueil dans telle

situation, et enfin construire et mettre en œuvre le recueil sont les compétences attendues chez un épidémiologiste. Les scénarios pédagogiques utilisés dans le cadre de la formation proposée à l'ISPED-Université de Bordeaux en épidémiologie visent ces compétences, et les stratégies pédagogiques permettent de mettre à l'épreuve les acquis des étudiants. Dans le cadre professionnel, le diplômé devra toujours se confronter à de nouvelles situations, ce qui justifie le choix de diversifier tant que faire ce peut les approches pédagogiques pour développer l'appétence à la construction appropriée de recueil de données.

Références

- [1] Ackoff, R. L. (1989), From data to wisdom, *Journal of applied systems analysis*, **16**(1), 3-9.
- [2] Franklin, C., G. Kader, D. Mewborn, J. Moreno, R. Peck, M. Perry and R. Scheaffer (2005), *Guidelines for assessment and instruction in statistics education (GAISE) report*, American Statistical Association, Alexandria, VA. Consulté à l'adresse http://www.amstat.org/asa/files/pdfs/GAISE/GAISEPreK-12_Full.pdf
- [3] Gattuso, L. (2011), L'enseignement de la statistique : où, quand, comment, pourquoi pas ?, *Statistique et Enseignement*, **2**(1), 5-30.
- [4] Jenicek, M. et R. Cléroux (1982), *Epidémiologie : principes, techniques, applications*, Edisem, Saint-Hyacinthe (Canada), Canada, France, Pays multiples.
- [5] Jutand, M.-A., A. Fagot Campagna et J.-P. Ferley (2012), *Chapitre 23 – Recueil des données épidémiologiques lors d'une enquête sur le terrain*. In: Dabis F. et Desenclos J.-C. (Eds), *Epidémiologie de terrain : Méthodes et Applications*, John Libbey Eurotext, Montrouge, pp. 191-205.