

Construire une enquête et un questionnaire

Hubert Maisonneuve, chef de clinique en MG, UFR Strasbourg
Jean-Pascal Fournier, chef de clinique en MG, UFR Toulouse

Auteur responsable de la correspondance

Jean-Pascal Fournier, Département universitaire de Médecine Générale, Faculté de Médecine de Toulouse Rangueil, Université Paul Sabatier, 133 Route de Narbonne, 31062 TOULOUSE Cedex, 05 62 88 90 48.

Les auteurs n'ont aucun conflit d'intérêt à déclarer.

Cet article a été rédigé à partir de la masterclass de Thomas Renaud (statisticien économètre) lors de la préconférence 2011 de FAYR-GP. FAYR-GP est l'association française des jeunes chercheurs en médecine générale. Créée en 2008, cette association loi 1901 a pour unique objectif de promouvoir la recherche en médecine générale auprès des internes et jeunes MG thésés depuis moins de 5 ans.

Retrouvez toutes les activités de l'association sur www.fayrgp.org.

Contexte

La méthodologie d'enquête et l'élaboration de questionnaires est issue du marketing. Elle repose pour beaucoup sur l'expérience et l'expertise. Dans cette « science molle » l'empirisme et le bon sens sont primordiaux. Il existe des ouvrages simples abordant sa méthodologie ^[1,2]. Internet regorge également de guides. Certains sont bien faits, d'autres moins. Il convient comme toujours, de rester critique face à un contenu non validé.

Cette méthode est fréquemment utilisée en recherche en médecine générale ^[3]. Il nous a paru pertinent d'en recadrer l'utilisation et les méthodes.

1. Concevoir son enquête en amont pour augmenter la compliance

L'enjeu final de l'enquête est de représenter fidèlement une réalité. Il est donc nécessaire de travailler sur :

- la participation des enquêtés à l'enquête,
- la compréhension des enquêtés. Elle concerne aussi bien les questions soulevées que l'enjeu de l'enquête.
- l'exhaustivité et la fiabilité des réponses. Toutes les questions doivent être conçue afin d'être répondues de manière cohérente, honnête et interprétable.

Ces trois notions se regroupent dans le concept plus global de compliance de l'enquêté, c'est-à-dire sa capacité à adhérer à l'enquête. La première difficulté de l'enquêteur est sous-tendue par la limite de compliance de la population étudiée. Plusieurs dimensions la mettent en péril et notamment celles qui rendent les questions de l'enquête compliquées ou peu compréhensibles (ex : personne âgée, problématique cognitive, appel aux souvenirs).

Il faudra donc prendre en compte des dimensions **cognitives et intellectuelles** mais également des dimensions **comportementales et sociales**. L'enquêté éprouve le besoin de se sentir dans la norme par rapport à l'idée qu'il se fait des attentes de l'investigateur (ex: enquête auprès de médecin sur leurs prescription d'arrêts de travail: vont-ils répondre sur leur pratique ou sur une idée qu'ils ont de la pratique attendue par l'enquêteur ?)

Pour maximiser la compliance, il faudra assurer un niveau de qualité important dans toutes les étapes de l'enquête :

- la conception du questionnaire (fond et forme)
- le choix du mode d'administration
- la sélection des enquêtés
- les «à-côtés» du questionnaire et du recueil
- les bénéfices secondaires à la participation (incitations et contreparties éventuelles).

2. Conception du questionnaire: grands principes

Cinq types d'informations peuvent être explorées dans un questionnaire

Opinion, satisfaction, croyance: il existe pour ces données un risque de biais de réponse en termes de désirabilité sociale. Attention donc aux questions «sensibles».

Autodescription: l'enquêté doit se décrire. Utiliser au maximum les questionnaires déjà établis.

Connaissances,

Comportements et pratiques,

Projections, intentions.

Différents types de questions

Question fermées/questions ouvertes

Aux questions fermées sont associées des propositions de réponses (par opposition aux questions ouvertes, où la réponse de l'enquêté est libre). On **privilégiera** autant que possible des questions **fermées**. Leur utilisation a pour conséquence une perte de précision, d'information et de nuance. Elles permettent néanmoins un recueil de réponses plus facile et plus rapides. Elles augmentent le nombre de répondants et simplifient l'analyse des données. En termes de méthodologie d'enquête, l'utilisation de questions fermées est nettement plus aisée et donne un comparateur fiable.

Dans le cadre d'enquête quantitative, l'utilisation de questions ouvertes doit rester **marginale**. On ne dépassera pas 20 à 25 % de questions ouvertes. Leur intérêt réside dans l'obtention de **précisions** en complément d'une question fermée. Elles permettent aussi de donner à l'enquêté l'impression qu'on lui donne la possibilité de s'exprimer. Elles améliorent son **adhésion** au questionnaire. Elles peuvent être utilisées uniquement à cette fin : les données ne seront alors pas interprétées.

Questions catégorielles (ou de liste)

Non ordonnées

Il n'existe pas d'ordre naturel entre les items (ex : catégorie professionnelle).

Généralement, ces questions sont **simples** et faciles à remplir. On peut donc les mettre au début du questionnaire pour **mettre l'enquêté en confiance**. Une autre option est de les placer à la fin car elles seront facilement remplies par un enquêté un peu «lassé» par le questionnaire. Ceci est essentiellement valable en cas de recueil auto-administré. En entretien, on mettra en confiance le questionnaire en les plaçant au début.

Si les items ne sont pas ordonnés, on gardera en tête que les catégories doivent toujours être mutuellement **exclusives**. L'utilisation de question à réponses multiples doit rester parcimonieuse. Les questions avec une seule réponse possible font gagner du temps et simplifient l'analyse. On les privilégiera, surtout en début ou en fin de questionnaire.

Lorsque plusieurs réponses sont possibles, on pourra proposer à l'enquêté de:

- cocher les réponses,
- les hiérarchiser en leur attribuant une note,
- cocher les réponses et choisir celle qui a le plus d'importance.

Ordonnées: questions d'échelle

On pourra utiliser selon les objectifs:

- des échelles de fréquence ou de notation (ex: de 1 à 10),
- des classes par «tranches» pour les variables continues (ex : 10-20 ans). Ceci limite la précision, mais à permet à l'enquêté qui n'est pas capable de donner une réponse précise et fiable, de se placer sa réponse dans un intervalle.
- des échelles de satisfaction ou de jugement.
ex: échelle de Likert: La personne interrogée exprime son degré d'accord ou de désaccord vis-à-vis d'une affirmation.
(ex: 1. pas du tout d'accord, 2. pas d'accord, 3. ni en désaccord ni d'accord, 4. d'accord, 5. tout à fait d'accord)

Les pièges à éviter

Attention aux réponses attractives!

Le questionné est attiré par les **premières** réponses de la liste (la première catégorie s'il y en a plusieurs). Il est également attiré par la modalité **médiane** (ex: sur une échelle de Likert: "ni en désaccord ni en accord"). Enfin, le questionné préférera les chiffres **ronds** dans les mesures quantitatives. On utilisera donc des classes par tranches plutôt que d'essayer d'obtenir le chiffre précis
ex: chiffrer une consommation de cigarette: le questionné ne va pas pouvoir chiffrer exactement sa consommation, il mettra facilement 10 ou 20. On pourra suggérer plus de précision en proposant des tranches «entre 10 et 15», «entre 15 et 20», etc.

Attention au bon nombre d'items!

Si la liste d'items est trop longue, on ne se souvient plus à la fin de ce qui avait été énoncé au début. Ceci est d'autant plus vrai pour les questionnaires téléphoniques. Si on en utilise trop peu, la précision de description du phénomène sera insuffisante. Le test du questionnaire avant diffusion permettra de s'approcher du bon nombre.

Attention à équilibrer les items!

Les items ne doivent pas être trop déséquilibrés entre eux.

ex : si 90 % des gens se retrouvent dans une catégorie, il n'y aura aucune possibilité d'expliquer la variabilité d'un phénomène entre les enquêtés. Il sera également intéressant de proposer une modalité « **autre** », afin de pouvoir identifier un phénomène non anticipé lors de la réalisation du questionnaire.

Attention à l'effet de liste!

L'utilisation de questions fermées se confronte à un risque d'effet de liste avec atomisation des réponses, lié notamment aux **choix terminologiques** (ex: quel choix de classification de maladies et jusqu'où aller dans la précision de la classification?). On effectuera donc nécessairement des regroupements. L'agrégation permettra d'augmenter la standardisation, mais entraînera un phénomène d'induction de la réponse. De même, l'utilisation de question ouverte entraînera également une **atomisation** des réponses mais le regroupement sera plus difficile.

Attention aux nombres de modalités dans une question à réponse sous forme d'échelle!

Une question récurrente est de déterminer dans une question à réponse sous forme d'échelle si l'on propose un nombre pair ou impair de modalités de réponses. Autrement dit, proposera-t-on une modalité neutre ou pas? Tout est fonction de l'objectif de l'enquête: on cherchera à **polariser** (=à mettre un nombre pair d'items) lorsque l'on voudra mener une enquête d'opinion, de satisfaction, quand on aura besoin que les gens **prennent parti**.

On pourra proposer une modalité médiane (ou une réponse type «sans opinion») pour des notions plus objectives (moins normatives, plus descriptives) comme les comportements, ou l'état de santé.



Quelques règles concrètes pour la rédaction du questionnaire

Leur respect doit permettre de limiter l'effort nécessaire à l'enquête pour répondre.

1- Formulation

Elle doit rester la plus simple possible. On évitera les concepts et les termes trop techniques, flous, les abréviations, les mots à double sens. L'utilisation de négation est à bannir (ex: "ne croyez-vous pas...").

2- Rester neutre

On prêtera attention à la tendance psychologique classique du questionné qu'est l'**acquiescement**. On évitera de laisser apparaître ses propres convictions ou d'utiliser des tournures naïves et consensuelles (ex "êtes-vous d'accord pour dire que..."). Ces dernières ont tendance à induire des réponses.

3- Utiliser des questions filtres

Il s'agit de questions qui conditionnent l'accès aux questions suivantes en fonction de la réponse donnée. Elles permettent d'analyser plus en détail quelques phénomènes rares ou spécifiques dans une population générale. On réserve ces questions spécifiques aux enquêtés ayant répondu positivement à la question filtre. Elles permettent d'aérer le questionnaire et à l'enquêté de ne pas lire toutes les questions. Elles diminuent le temps de passation du questionnaire.

4- Structurer le questionnaire

On constituera des blocs **logiques** regroupés par thèmes. Ils améliorent la compréhension globale par l'enquêté.

La taille et la forme du questionnaire sont primordiales pour s'assurer d'un maximum de retour: il doit être court, lisible et aéré. Enfin, il est indispensable de numéroter les questions!

5- Soigner son accroche

La première question est importante. Il s'agit le plus souvent d'une question «brise-glace», qui se doit d'**accrocher** l'enquêté.

Identifier et anticiper les biais

Limiter l'effort de mémoire (notamment dans le cas de recueil rétrospectif sur période longue)

Limiter la durée d'informations requises, par exemple à un mois ou à une semaine permettra de fiabiliser la réponse.

Limiter le risque de désirabilité sociale

Ce risque dépend des pays et des thèmes explorés par l'enquête : argent, santé, comportement sexuel... On choisira attentivement le mode d'administration du questionnaire. On utilisera une formulation empathique sur les questions sensibles. Ces notions sont également primordiales lorsque l'on s'attache à mesurer des connaissances ou des compétences. Par exemple, à la question « connaissez-vous le référentiel X ? » L'enquêté aura tendance à répondre oui, et ce même si le dit référentiel n'existe pas.

S'inspirer de l'existant

Il est nécessaire d'explorer les données de la littérature «classique», des thèses, mais aussi des enquêtes publiques (démographie médicale, économie de la santé) afin d'utiliser des éléments de questionnaires déjà utilisés et validés.

ex: nomenclatures INSSEE pour les profils socio-démographiques, minimodule européen ou SF36 pour décrire des états de santé, etc.

3. La mise en œuvre de l'enquête

Les «à cotés» du questionnaire

L'enquêté doit être **prévenu** qu'il va être enquêté. Ceci permettra de lui expliquer les points importants pour la bonne marche de l'enquête. Cette information peut être donnée sous la forme d'un courrier accompagnant le questionnaire, ou mieux d'un pré-courrier envoyé à l'avance. Le **mode d'envoi** et le **délai** seront définis dans le protocole. Le **pré-courrier** permet de donner le cadre de réalisation de l'enquête, de présenter l'objet d'étude, d'aider à sensibiliser l'enquêté au projet et donc de maximiser sa compliance. Il permettra à l'enquêteur d'asseoir sa **légitimité** auprès de l'enquêté. Il permet aussi de le rassurer sur le respect des aspects **réglementaires**, légaux et éthiques (ex: questions d'anonymat, de conservation des données, de déclaration à la CNIL). Le **temps de passation** du questionnaire est à annoncer impérativement. Enfin, ne pas oublier les formules de politesse habituelles.

On définira d'avance le délai entre information et enquête. On peut prévoir un certain nombre de relance afin d'augmenter la population d'étude. La forme des relances visera à remémorer l'étude sans braquer les éventuels participants.

Un pré-test systématique du questionnaire

Il aura pour but de vérifier :

- la conception des questions,
- l'adaptation des modalités de question,
- la présence d'éventuels oublis,
- d'évaluer le temps réel de passation.

Ce pré-test **systématique** nécessite d'être réalisé dans des conditions **similaires** à celles dans lesquelles le questionnaire final sera renseigné: même mode d'administration et enquêtés similaires. Afin d'éviter tout ce qui pourrait être implicite, on testera des personnes ne connaissant pas le sujet. L'échantillon du pré-test dépend de l'ampleur de l'enquête. D'une manière générale, on considère qu'environ **1 à 10 %** de l'effectif global est une proportion acceptable.

La collecte des données

Quand : quelle est la période la mieux choisie pour étudier le phénomène? Il faut définir d'avance le planning, avec un échéancier et des possibilités pour rattraper un éventuel retard (ex: éviter de lancer le questionnaire en pleine période de vacances). Si nécessaire, des relances seront prévues à différentes périodes.

Comment : quelles sont les modalités de passation et de récupération du questionnaire?

Face à face : il permet d'obtenir les informations les plus détaillées: la qualité et la complexité des infos recueillies sont meilleures. Dans une certaine mesure, il permet de gommer les pièges ou les erreurs du questionnaire. Ce mode a tendance à **augmenter** l'adhésion et la **compliance** globale, dans les limites du relationnel de l'enquêteur. Il pose le problème de l'investissement temporel et financier. Il peut être difficile de joindre les enquêtés. Il peut exister un **biais systématique entre les enquêteurs** ou une majoration du risque de **désirabilité** sociale (il est plus difficile pour l'enquêté de prendre une position tranchée et/ou mal vue en face à face).

Téléphone : c'est un mode peu coûteux. Il est très intéressant pour un questionnaire **court** (attention au risque de lasser l'enquêté si le questionnaire est trop long). Pour cette même raison, il peut être utile de réunir les questions intéressantes au début du questionnaire. Les questions sont nécessairement simples, bien structurées, avec peu de propositions de réponses. Ce mode de passation est intéressant lorsqu'on pose des questions sensibles (plus grande

neutralité de l'enquêté vis-à-vis de l'enquêteur). La difficulté pour joindre les sujets reste élevée (où et quand joindre les enquêtés?).

Questionnaire auto-administré (courrier ou mail) : c'est la méthode la plus souvent utilisée, car très peu coûteuse. Elle renforce le sentiment d'anonymat de l'information transmise, et permet à l'enquêté de répondre au moment qu'il juge le plus adéquat. Ces questionnaires nécessitent d'être courts et clairs: l'absence d'interaction et de réajustement oblige à donner des **explications limpides**.

Pour les questionnaires électroniques, il existe de plus en plus de logiciels permettant de les construire, de les éditer, de les mettre en ligne (ex: LimeSurvey, Google docs, Qualtrics...). Les questionnaires internet, comme les questionnaires envoyés par courrier, souffrent d'un nombre important de **non réponse**. On répondra partiellement à cette limite en choisissant attentivement une population d'étude concernée par la question posée.

Qui cible-t-on? Qui notre enquête a-t-elle vocation à représenter?

Les réponses à ces questions sont données dès la rédaction du protocole. La question de recherche ^[4] à laquelle l'enquête va répondre définit le champ exploré. Comme dans toute étude quantitative, cela conditionnera les critères d'inclusion, d'exclusion, et de représentativité à mettre en œuvre.

La **représentativité** d'un échantillon sous-entend que celui-ci est un «modèle réduit» représentatif de la population à étudier. Les critères de représentativité (âge, sexe, zone géographique...) seront définis à l'avance. Les conclusions obtenues sur ce modèle seront alors extrapolables à l'ensemble de la population à laquelle on s'intéresse. La représentativité peut être visée à priori grâce au mode de sélection de l'échantillon, ou à posteriori, à l'aide de critères correctifs.

Pour les enquêtes par questionnaire, le recrutement par **volontariat** paraît obligatoire. Néanmoins, à la longue, les éternels volontaires **s'usent**, portent moins d'attention aux réponses et causent une déperdition d'information, ou pire des réponses non fiables. On peut citer l'exemple des maîtres de stages en médecine générale qui sont soumis à un très grand nombre de propositions. Une solution peut être alors de baisser le nombre de participants à une enquête afin d'augmenter le sentiment d'implication et favoriser le taux de réponse. Afin d'éviter le volontariat exclusif, on doit systématiquement explorer les **incitatifs et contreparties** (récompense financière par exemple). La **contrepartie** augmente la participation et l'adhésion, mais cause un **biais de sélection**. Lorsque la participation est indemnisée d'une manière ou d'une autre, il existe un risque **professionnalisation** des répondants: ces

participants s'inscrivent systématiquement, ce qui finit par nuire à la fiabilité. Il faut envisager systématiquement les contreparties non monétaires: il faut savoir impliquer l'enquêté dans les conséquences du résultat. Prévoir d'avance une publication liée à l'enquête et le mentionner dans le descriptif envoyé implique plus l'enquêté qu'un mémoire ou une thèse. Dans ces conditions on lui proposera systématiquement de lui envoyer les résultats et/ou l'article écrit (**feedback**). On pourra éventuellement personnaliser les résultats, avec possibilité de situer les enquêtés les uns par rapport aux autres (méthode de **benchmarking**).

En cas de refus de participation, on demandera le motif. Si le rejet est majoritaire, il peut être intéressant de monter un protocole explorant les motifs, afin d'adapter les études ultérieures.

L'échantillonnage aléatoire nécessite d'avoir une base de donnée exhaustive afin de pouvoir randomiser et créer un échantillon initial important (voire très important) pour anticiper sur l'importance de la non réponse. Des bases de données existent dans le cas des médecins, il est utile de les connaître : ordre des médecins, CPAM, ADELI, RPPS, AMELI, URPS...

La méthode des quotas: on crée des critères de sélection nous permettant de retenir un échantillon à partir d'une population globale, sans caractéristique particulière. On appelle les gens au hasard, et on ne garde que les sujets se conformant au profil. Il est nécessaire de bien connaître les caractéristiques de la population de référence afin d'en dégager de bons critères de représentativité. Il persiste un risque de biais de l'échantillon si ces critères sont mal choisis.

Si l'on compare les deux méthodes, le sondage aléatoire est plus adapté pour les échantillons de plus de 100 personnes, la méthode des quotas correspond assez bien à de petits échantillons.

La méthode du pas systématique: on sélectionne le sujet en fonction d'un pas qui sera toujours le même sur un critère choisi d'avance (ex : on sélectionne un participant tous les x consultations, ou toutes les x minutes). Cette alternative présente comme intérêt de ne pas nécessiter de fichier de référence.

L'un des défauts est la **possible non réponse**. Plaçons-nous dans le cas où on sélectionne tous les x. On ne sait pas si tout le monde va répondre. Il est donc nécessaire de prendre un pas suffisamment rapproché.

Si l'un des participants ne répond pas, on laisse un pas de x avant d'interroger le suivant.

ex : on dispose d'une population cible de 240 médecins, on souhaite en interroger vingt, soit un sur douze. On prévoit un refus de

participation de trois médecins sur quatre. On peut choisir un pas de 3 et proposer la participation à un médecin sur trois.

Taille minimale ou optimale de l'échantillon.

Le minimum «absolu» se situe entre 20 et 30 répondants aléatoires, sauf dans le cas de population particulièrement petite et ciblée.

La taille optimale d'un échantillon dépend en réalité de la méthode de sélection de l'échantillon retenu (aléatoire vs. quota), de la fréquence du phénomène qu'on veut observer dans la population, de la précision souhaitée, et enfin, de la taille de la population de référence.

La réelle problématique est d'anticiper le taux de non répondants, pour s'avoir à combien de sujets potentiels adresser le questionnaire et s'assurer d'un nombre de réponse suffisant.

Conclusion

L'enquête par questionnaire est un outil intéressant pour la recherche quantitative en médecine générale. Cette méthode répond à de nombreuses règles, ayant pour objectif de maximiser la compliance de l'enquêté. Nous vous proposons une checklist des principaux points à ne pas oublier.

Références

1. Fenneteau H. Enquête: entretien et questionnaire. 2e éd. Dunod; 2007.
2. De Singly F. Le questionnaire: L'enquête et ses méthodes. 3e éd. Armand Colin; 2012.
3. Kruschinski C, Lange M, Lionis C, van Weel C, Hummers-Pradier E. Themes and methods of research presented at European General Practice Research Network conferences. *Fam Pract.* 2010;27(4):459-7.
4. Centre Cochrane Français. Tutoriel «Introduction à l'Evidence-based nursing»
1. Formuler la question [En ligne]. 2011 [Consulté le 14/08/2012] Consultable à l'URL: <http://bit.ly/V1aagg>

Check-list enquêtes par questionnaires

Check-list protocole

- Le mode de distribution du questionnaire est le plus adapté à l'enquête.
- Un nombre adapté de relances est prévu.
- Le questionnaire prévu utilise des éléments de questionnaires déjà utilisés et validés.
- Le questionnaire a été testé sur un nombre suffisant et représentatif de participants.
- La population d'étude est adaptée à la question de recherche.
- Le questionnaire concerne un échantillon représentatif de la population d'étude.
Le mode d'échantillonnage est adapté à la question de recherche.

Check-list lettre d'accompagnement

- La formulation de la lettre est adaptée et les formules de politesses sont présentes.
- Les coordonnées de l'enquêteur et de son directeur sont précisées.
- L'objectif de l'enquête est expliqué.
- Le temps de passation est précisé.
- Le nombre de questions est annoncé.
- L'anonymat des réponses est précisé.
- Un retour sur les résultats de l'étude est envisagé pour les participants.
Les modalités en sont précisées (mémoire, thèse, publication).

Check-list questionnaire

Ergonomie du questionnaire

- Les questions sont toutes numérotées.
- L'objectif de l'enquête est rappelé sur le questionnaire.
- Les premières questions sont simples et accrocheuses.
- Le questionnaire est organisé par blocs logiques, du général au particulier.
- Il existe des questions de transition entre les blocs.
- L'ordre des blocs peut être justifié par l'enquêteur.
- Les questions ne débordent pas sur la page suivante.

Forme des questions

- Chaque question correspond à une idée unique.
- Les questions sont formulées de façons précises et concises.
- Les termes utilisés dans le questionnaire ont été adaptés à la population de l'enquête.
- Les formulations des questions sont neutres.
- L'utilisation d'éventuelles questions ouvertes répond aux objectifs de l'enquête.

Forme des réponses

- Le type de réponses est adapté à la question de recherche (ouverte/fermée, dichotomique vs choix multiple).
- Le choix du support est adapté aux critères mesurés (échelle de Lickert vs item).
- Le nombre d'items est adapté à l'objectif.