



**HAL**  
open science

## **Concevoir, traduire et valider un questionnaire. A propos d'un exemple, EUROQUEST.**

A. Bouletreau, D. Chouaniere, P. Wild, J.M. Fontana

### **► To cite this version:**

A. Bouletreau, D. Chouaniere, P. Wild, J.M. Fontana. Concevoir, traduire et valider un questionnaire. A propos d'un exemple, EUROQUEST.. [Rapport de recherche] Notes scientifiques et techniques de l'INRS NS 178, Institut National de Recherche et de Sécurité (INRS). 1999, 46 p., ill., bibliogr. <hal-01420163>

**HAL Id: hal-01420163**

**<https://hal-lara.archives-ouvertes.fr/hal-01420163v1>**

Submitted on 20 Dec 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

MAI 1999

N° ISSN 0397 - 4529

# 178

## **Concevoir, traduire et valider un questionnaire A propos d'un exemple, EUROQUEST**

**Agnès Bouletreau, Dominique Chouanière,  
Pascal Wild et Jean-Marc Fontana.  
Service d'Epidémiologie, INRS**

**INSTITUT NATIONAL DE RECHERCHE ET DE SECURITE**

**SIEGE SOCIAL :  
30, RUE OLIVIER-NOYER, 75680 PARIS CEDEX 14**

**CENTRE DE RECHERCHE :  
AVENUE DE BOURGOGNE, 54501 VANDŒUVRE CEDEX**

# **Concevoir, traduire et valider un questionnaire A propos d'un exemple, EUROQUEST.**

**Agnès Bouletreau, Dominique Chouanière, Pascal Wild et Jean-Marc Fontana.  
Service d'Epidémiologie, INRS**

Les auteurs remercient vivement Marc Favaro, Unité « Accidentologie » - INRS, pour la relecture attentive de ce texte et ses conseils avisés.

Exemplaire à demander à D. Chouanière, Service Epidémiologie, INRS,  
Avenue de Bourgogne, BP n° 27, F - 54 501 Vandoeuvre, France.  
Tel : 03 83 50 21 22, fax : 03 83 50 20 15, E-mail : [chouaniere@inrs.fr](mailto:chouaniere@inrs.fr)

## RESUME

Les questionnaires sont des outils de recueil d'information souvent irremplaçables en recherche. Les échelles multidimensionnelles sont les mieux adaptées à l'exploration des maladies ou dysfonctionnements intégrant plusieurs composantes (telles que la dépression ou l'altération de la qualité de vie). Ces échelles mises au point par les psychosociologues, sont, du fait de l'ouverture du champ des études biomédicales vers les domaines psychosociaux, de plus en plus employées.

Constatant que le transfert de connaissances et d'expérience, quant à l'élaboration et l'évaluation d'un questionnaire, entre les disciplines psychosociologiques et biomédicales, est peu développé, nous avons décidé de confronter les deux disciplines. Cette tentative de synthèse s'est avérée riche, elle a permis de définir les principes de rédaction d'un questionnaire ainsi qu'une méthodologie de traduction garantissant, quelle que soit la langue, la permanence des concepts ; elle a rendu possible la comparaison des méthodes d'évaluation tant qualitatives que quantitatives. Si les auteurs sont unanimes sur la nécessité de mesurer la fiabilité, la sensibilité au changement et la validité d'un questionnaire, il n'y a pas, par contre, de consensus sur le concept de validité et des méthodes statistiques pour l'explorer.

La mise à l'épreuve de ces méthodes sur EUROQUEST (questionnaire de symptômes neurotoxiques pour des salariés exposés à des produits chimiques ayant des effets sur le système nerveux) a mis en lumière l'inadéquation, pour cet exemple, de certaines méthodes classiquement décrites. Le calcul des odds ratios entre deux items ou la classification hiérarchique ascendante se sont, par contre, avérées être des méthodes fructueuses.

**MOTS-CLES** : Validation, Questionnaire, EUROQUEST, Neurotoxicité, Méthodes statistiques, Fiabilité, Sensibilité au changement, Validité

# SOMMAIRE

## 1. Introduction

## 2. Elaboration du questionnaire

- 2.1. Conception du questionnaire
  - 2.1.1. Elaboration des questions
  - 2.1.2. Les différentes modalités de réponse
    - \* Réponses en clair
    - \* Echelles visuelles analogiques
- 2.2. Evaluation qualitative
- 2.3. Traduction d'un questionnaire

## 3. Validation d'un questionnaire

- 3.1. « Scoring » : construction d'un score
- 3.2. Evaluation quantitative d'un questionnaire
  - 3.2.1. Etude de la fiabilité
    - \* Fiabilité liée à l'enquêteur
    - \* Fiabilité au cours du temps
  - 3.2.2. Etude de la validité
    - \* Validité d'apparence
    - \* Validité de contenu
    - \* Validité contre critère
    - \* Validité de structure
      - validité de structure externe
        - . validité convergente
        - . validité divergente
      - validité de structure interne
        - . méthode des odds-ratios
        - . classification hiérarchique ascendante
  - 3.2.3. Etude de la sensibilité au changement

## 4. Discussion

## 5. Conclusion

Bibliographie

Tableaux

Annexe

## 1. INTRODUCTION

Les questionnaires sont de plus en plus utilisés en recherche biomédicale. Leur rôle dans le dispositif du recueil des informations est variable. Couplés à d'autres mesures, ils peuvent n'être que de simples outils complémentaires, tel est le cas des questionnaires de symptômes respiratoires souvent associés à une exploration fonctionnelle en épidémiologie pulmonaire. Mais les questionnaires sont parfois le seul moyen de recueillir des données, soit qu'il s'agisse de faits passés tels que les antécédents médicaux ou l'exposition professionnelle à un produit toxique, soit qu'il s'agisse d'évaluation de connaissances en matière de contraception, par exemple, dans la perspective d'une action d'éducation à la santé ; soit encore qu'il s'agisse de perception, de comportement ou d'opinion ; l'évaluation d'une chimiothérapie peut comporter un volet d'appréciation, par les malades, de la "qualité de vie" mesurée par questionnaire.

A cette classification basée sur la nature de l'information à recueillir, s'ajoute une typologie selon la normalisation des questions. Certains questionnaires explorent des maladies dont la définition fait l'objet d'un consensus comme la bronchite chronique. Dans ce cas, des questions normalisées permettent d'établir le diagnostic : *Toussez-vous presque tous les jours pendant 3 mois chaque année ? Etes-vous essoufflé(e) au repos ?* (questionnaire international des affections respiratoires) [30]. D'autres questionnaires étudient des maladies ou dysfonctionnements avec plusieurs facettes ou dimensions. Chaque composante de la maladie est explorée par plusieurs questions formant un domaine particulier du questionnaire. Dans ce cas les questions sont rarement normalisées. On appelle ce type de questionnaire, les "échelles multidimensionnelles" [10]. Les questionnaires d'évaluation de la maladie dépressive en sont un exemple, comportant généralement les domaines tels que "la perte des sensations émotionnelles", "le ralentissement psychomoteur et intellectuel", "les troubles du sommeil", "l'altération de l'humeur", etc.

Depuis quelques années, ces échelles ont largement investi le domaine biomédical. Ce développement s'explique d'une part, par l'ouverture du champ biomédical, en particulier épidémiologique, à des thèmes d'étude psychosociaux tels que la qualité de vie, le handicap, etc. et d'autre part par l'évolution même du concept de maladie ; à la conception pastorienne basée sur l'effet d'un seul agent sur un organisme, s'est imposée la conception plurifactorielle de la maladie avec l'inclusion des facteurs psychosociaux comme agents promoteurs des altérations de la santé [30, 35].

Une attention particulière doit être portée à la validation des échelles multidimensionnelles, à structure complexe, dont il sera particulièrement question ici. Comme pour tout instrument de mesure, les qualités du questionnaire doivent être connues avant de débiter une étude. On ne pourra prendre en compte les données fournies par le questionnaire qu'après avoir vérifié trois de ses qualités fondamentales : les mesures réalisées sont bien reproductibles, l'outil choisi mesure bien ce pour quoi il a été construit et il permet de rendre compte d'une modification même minime du phénomène exploré, au sein d'un groupe ou chez un même individu. En d'autres termes, il s'agit d'apprécier la fiabilité, la validité et la sensibilité au changement du questionnaire. Une fiabilité acceptable permet de minimiser les erreurs aléatoires de mesure et une validité correcte permet de réduire les erreurs systématiques de mesure. La sensibilité au changement du questionnaire doit être connue afin de choisir un instrument adapté au phénomène mesuré, de la même manière que la connaissance de la sensibilité du pèse-personne ou de la balance de laboratoire permet de les récuser lorsque l'objet à peser avoisine la tonne. [10, 17, 18, 31, 35].

Quand le thème d'étude est entièrement neuf, les équipes peuvent se lancer dans l'élaboration d'un questionnaire. Cette étape peut, à elle seule, constituer un travail de recherche [18, 31].

A mi-chemin entre l'élaboration d'un nouveau questionnaire et la validation d'un outil existant, se trouve la traduction d'un questionnaire, validé dans une autre langue. Pour garantir la comparabilité entre la version originale et la version traduite, il est recommandé de se conformer à une procédure de traduction précise [21].

Depuis longtemps les sciences sociales ont utilisé le questionnaire comme outil de recueil [2, 4, 13, 19, 33]. De ce fait, les étapes précédemment envisagées, de la conception à la validation, ainsi que la méthodologie à employer ont été largement décrites dans la littérature psychosociale.

Mais il nous a semblé que l'expérience capitalisée, par les disciplines psychosociales notamment, dans l'élaboration des questionnaires était largement sous-utilisée en épidémiologie.

Par ailleurs, le transfert des méthodes d'évaluation des questionnaires des disciplines psychosociales au domaine biomédical ne va pas sans poser des questions. En effet, les objectifs de ces deux domaines de recherche sont parfois opposés. Si les études psychosociales cherchent à décrire chaque nuance d'un comportement, les épidémiologistes ont tendance à considérer plutôt la structure psychologique des individus comme un éventuel facteur de confusion et à essayer de la gommer afin d'en affranchir les résultats [35]. De plus, l'objet même des disciplines est différent : les sciences sociales décrivent volontiers les facettes d'un problème, par exemple celles de l'échec scolaire, sans forcément s'intéresser aux relations causales entre le problème étudié et les facteurs identifiés. L'épidémiologie analytique, au contraire, essaie d'accumuler les preuves en faveur de cette relation de causalité, comme, par exemple, la recherche de la relation entre prématurité et échec scolaire.

Ainsi les objectifs qui ont présidé à cette réflexion et à cette étude sont pluriels :

- inventorier les connaissances les plus communément admises dans les sciences psychosociales et en épidémiologie en ce qui concerne la méthodologie d'élaboration, de traduction et d'évaluation des questionnaires,
- analyser les différences, les possibilités de transfert entre disciplines,
- proposer une méthodologie d'évaluation des questionnaires, pertinente pour, et applicable à l'épidémiologie,
- appliquer cette méthode d'évaluation à titre d'exemple, à un questionnaire d'évaluation des symptômes neurotoxiques, EUROQUEST,
- discuter, à la lumière de ces premiers résultats, des axes de progrès possibles.

Dans une première partie, seront rappelées, les règles de rédaction et de traduction d'un questionnaire. Puis, les méthodes de validation, étude de la fiabilité, de la validité et de la sensibilité au changement, seront présentées selon le point de vue des disciplines biomédicales et psychosociologiques et appliquées à EUROQUEST [10, 12, 17, 31]. Le dernier chapitre de discussion tentera de tirer les enseignements de ces premiers résultats.

## 2. ELABORATION D'UN QUESTIONNAIRE

Ce chapitre est consacré aux trois étapes évoquées plus haut : élaboration, traduction et validation. L'exemple dont il sera question, est l'EUROQUEST (European Questionnaire). Cet autoquestionnaire a été élaboré en anglais en 1992 par le réseau EURONEST (EUROpean NEurological Solvent Toxicity) afin de détecter les symptômes neurotoxiques chez des salariés exposés professionnellement et de façon chronique à certains agents chimiques [21]. Il a été utilisé dans deux études :

- une étude longitudinale prospective portant sur les effets d'une exposition chronique aux solvants contenus dans les peintures ; depuis 1993, sont suivis 408 apprentis dont 201 peintres et 207 "témoins" apprentis dans un autre métier du bâtiment.
- une étude transversale sur les effets neurotoxiques du toluène. Cette étude a été conduite auprès de 128 salariés de deux imprimeries.

La validation d'EUROQUEST porte sur ces 536 questionnaires.

### 2.1. CONCEPTION D'UN QUESTIONNAIRE

Cette étape comprend l'élaboration des questions, des choix de réponses ainsi que le pré-test [31]. Peu d'articles méthodologiques portant sur les questionnaires et appartenant au champ de l'épidémiologie s'intéressent à cette étape ou sa suivante qui concerne l'évaluation qualitative. Les remarques qui suivent sont essentiellement tirées de la littérature sociologique.

#### 2.1.1. ELABORATION DES QUESTIONS

La rédaction d'un questionnaire doit être précédée d'une définition des hypothèses générales de l'enquête ainsi que de ses objectifs [13, 33]. Ils contribueront à définir les aspects du problème qui seront recueillis par le questionnaire et à préciser l'information désirée [2, 5]. Ces informations seront explorées par des "items"; un "item" est la plus petite unité d'information pouvant être isolée par une mesure ou un examen. Ces "items" peuvent être regroupés en "domaines", quand les problèmes ont plusieurs composantes telles celles de l'état dépressif mentionnées dans l'introduction.

Lors de la rédaction, certains points doivent être considérés avec attention. Le vocabulaire employé doit être compréhensible par toutes les personnes de la population cible et les questions doivent avoir le même sens pour tous [19, 20]. La rédaction d'un questionnaire peut être précédée d'une série d'entretiens non ou semi-directifs auprès d'un échantillon représentatif de la population cible. Ils permettent de préciser l'information à recueillir et ses modalités d'expression. L'analyse sémantique du vocabulaire utilisé par les enquêtés permettra de retenir dans le questionnaire les mots clés et les énoncés les plus employés [31].

Un questionnaire ne doit pas être considéré comme une juxtaposition de questions indépendantes. Il doit être construit comme un entretien et planifié comme un tout. Son acceptabilité en dépend. Les questions doivent se suivre dans un ordre logique non seulement pour que la personne interrogée prenne plaisir à y répondre mais aussi pour qu'elle se construise une représentation du phénomène étudié. Les thèmes délicats, tels que les opinions politiques, ou les pratiques sexuelles ou religieuses, doivent être abordés avec précaution. Pour traiter de tels sujets, on a recours à des moyens tels que la méthode dite du "thème noyé"

où le point clé est dissimulé dans un contexte plus large ; par exemple, on ne demandera pas à un catholique s'il va à la messe, mais on lui demandera de décrire l'emploi du temps de ses trois derniers dimanches [33].

Un questionnaire n'est réellement informatif que si les réponses apportées sont sincères et réfléchies. Pour ce faire, il faut prêter une attention particulière à la présentation du questionnaire. En effet l'induction de réponses, attribuable à un libellé peu judicieux de l'item ou à un agencement particulier des questions, est fréquemment citée dans la littérature [33]. Nous avons relevé quelques unes des erreurs à éviter.

La "réaction de prestige" ou "tendance de façade" correspond à la peur de se faire mal juger à travers ses réponses. Elle s'applique aux questions commençant par des propositions telles que, *Selon vous*, ou, *A votre avis*. On remarque pour ce type de questions une plus grande hésitation à répondre se traduisant par une augmentation des réponses, *Je ne sais pas*, ou par des refus de répondre.

"L'attraction de la réponse positive" est une tendance connue depuis longtemps en psychologie sociale. Elle serait corrélée à la sensibilité aux placebos et serait donc une expression de la suggestibilité. Cette erreur systématique de réponse est amplifiée par un énoncé tendancieux ; on obtient alors des réponses totalement suggérées par la question. Pour une formulation de ce type, *Les qualités pédagogiques vous paraissent-elles essentielles pour un être bon professeur ? Oui - Non*, la réponse attendue/souhaitée par l'enquêteur apparaît clairement.

Enfin la peur de certains mots, ayant une charge affective et/ou socialement inacceptables, influence les réponses. Ainsi, en octobre 1941, dans un sondage d'opinion portant sur la participation des USA à la guerre, l'énoncé, *Faut-il déclarer la guerre*, jugé trop inquiétant après le pré-test, a été remplacé par, *Faut-il prendre part au conflit ?* [33].

Le questionnaire exige de la personne qui s'y soumet, qu'elle fasse un effort d'attention et de réflexion et éventuellement qu'elle "s'expose" si les questions touchent à des thèmes intimes. La présentation du questionnaire, en exacerbant ou au contraire en gommant ces difficultés, peut être à l'origine d'erreurs systématiques dans le recueil des données. Ainsi lorsque le questionnaire débute par des items trop "personnels" ou "délicats", la réaction de "contraction défensive" des enquêtés sera particulièrement fréquente et le nombre de questionnaires non remplis s'en trouvera artificiellement augmenté. Cette réaction de contraction défensive se manifeste également lors des changements trop brusques de thèmes à l'intérieur d'un questionnaire ; faute de questions de liaison entre deux domaines, les enquêtés peuvent être suffisamment déroutés pour abandonner le questionnaire.

L'effet "halo" ou "effet de contamination" est un autre mécanisme pouvant être responsable d'erreurs systématiques dans le recueil des données. Suite à un sentiment négatif, d'irritation, par exemple, provoqué par une question, les réponses aux questions suivantes vont être éventuellement faussées. La "contagion" peut aussi se produire par une autocensure chez des personnes voulant préserver la cohérence de leurs réponses à des questions proches, au détriment de la vérité. Par exemple, quelqu'un qui s'est déclaré "catholique pratiquant" sera tenté de répondre qu'il va tous les dimanches à la messe, alors qu'il ne fréquente l'église qu'un fois par mois.

Enfin, le concepteur d'un questionnaire doit garder en tête que la longueur des questions et/ou du questionnaire peut provoquer la "démission" de l'enquêté [11].

Les facteurs de personnalité qui exagèrent ou minimisent la déclaration des symptômes peuvent être pris en compte dans un questionnaire. L'inclusion d'une grille prenant en compte les hésitations, en dénombrant, par exemple, les réponses, *Je ne sais pas*, permet de mesurer indirectement le degré d'anxiété de la personne interrogée. L'émotivité ou la tendance

hypochondriaque peuvent être appréciées grâce à l'ajout dans le questionnaire, d'une échelle MMPI (Minnesota Multiphasis Personality Inventory). Dans le cas d'EUROQUEST, deux traits psychologiques pouvant influencer la déclaration des symptômes ont été considérés, l'anxiété et l'hypersensibilité à l'environnement.

La reproductibilité des réponses peut être testée en introduisant des questions "pièges" ou redondantes [33, 35]. Par exemple, la réponse à une première question, *Avez-vous déjà eu des crises d'asthme ?* peut être confirmée par une seconde question, rédigée de façon différente, située plus loin dans le questionnaire, *Avez-vous déjà eu des crises d'étouffement avec des sifflements dans la poitrine ?* [29].

### 2.1.2. LES DIFFERENTES MODALITES DE REPONSE

Lors de la rédaction d'un questionnaire, les modalités de réponse vont être un point important de la réflexion.

**Les réponses en clair :** quatre types de formulation sont possibles [35] :

**Les réponses "ouvertes" :** la personne répond aussi librement que lors d'un interview. Cette modalité tend à avantager ceux qui s'expriment facilement et qui ont un vocabulaire riche.

**Les réponses dichotomiques ;** les questions sont "fermées" ne laissant que les choix : *Oui* ou *Non*, *Vrai* ou *Faux*, etc. Ces réponses sont fréquemment employées car leur traitement informatique est facile et ils contraignent les hésitants à s'exprimer clairement.

**Les items trichotomiques :** variante des précédentes pour lesquelles le choix est élargi à la modalité, *Je ne sais pas*.

**Les réponses de reconnaissance ou "cafétéria" ou à choix multiple :** plusieurs réponses de nature qualitative sont proposées, comme dans l'exemple, *Etes-vous professionnellement exposé(e) à des :*

- . Réponse 1 : *Atmosphères humides*
- . Réponse 2 : *Atmosphères froides*
- . Réponse 3 : *Courants d'air*

**Les réponses qualitatives ordonnées :** les réponses proposées ont une gradation comme dans l'exemple : *Avez-vous mal à la tête ?*

- . Réponse 1 : *Jamais/Parfois (3 à 4 fois par an)*
- . Réponse 2 : *Souvent (au moins une fois par mois)*
- . Réponse 3 : *Tous les jours*

Dans ce type de réponses, la quantification est sous-jacente et le concepteur du questionnaire doit apprécier l'écart entre les différentes réponses ainsi que les équivalences entre ces écarts. Dans l'exemple précédent, il s'agit d'évaluer si l'écart entre la réponse 1 et 2 est le même que celui entre la réponse 2 et 3. Le traitement futur des données en découle directement. Les tests paramétriques ne peuvent être employés que si l'hypothèse de continuité : même "distance" entre chaque réponse possible, est retenue [10, 31]. Une autre solution est de transformer, lors de l'analyse, les réponses qualitatives ordonnées en réponses dichotomiques.

On parle de réponse unipolaire lorsque l'une des graduations extrêmes représente l'état "normal" ; dans l'exemple, *Je suis essoufflé(e) en marchant même lentement : Jamais, Parfois, Souvent*, la réponse, *Jamais*, représente l'état normal. On parle de réponse bipolaire lorsque l'état "normal" se situe entre les deux graduations extrêmes. Pour la

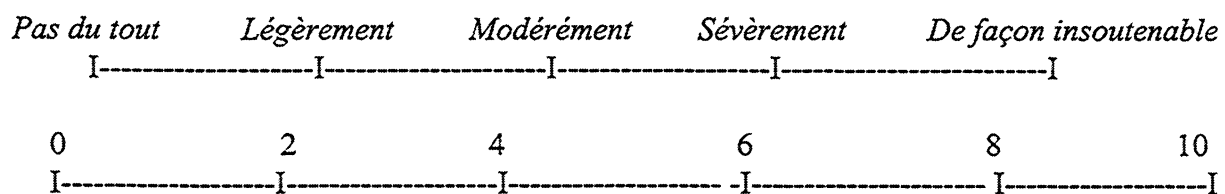
question, *Avez-vous du mal à vous endormir ? Jamais, Parfois, Souvent*, la réponse, *Parfois*, définit l'état normal.

Le nombre des possibilités de réponses doit être choisi avec discernement. Un nombre de propositions paires oblige à prendre position en supprimant la tendance, fréquente parmi les indécis, de choisir la note centrale. De même, en introduisant dans les réponses bipolaires des graduations volontairement excessives, comme, *Jamais* ou *Toujours*, on peut contourner la réticence habituelle à employer les propositions extrêmes [31].

D'autres modalités de réponses existent mais elles sont moins utilisées dans le domaine biomédical.

### Les échelles visuelles analogiques :

Elles sont souvent employées pour appréhender l'intensité ou la fréquence de certains symptômes. Ces échelles augmentent la sensibilité de la réponse puisque qu'elles élargissent l'éventail des choix ; on y a souvent recours pour mesurer les phénomènes douloureux ; à la question, *Avez-vous mal au poignet ?* les propositions de réponses peuvent se présenter ainsi :



Mais la fiabilité de ces échelles pose problème ; leur présentation peut, à elle seule, induire des réponses différentes. En effet, des études ont mis en évidence un manque de reproductibilité des réponses selon que celles-ci sont cochées sur des échelles orientées dans le sens horizontal ou dans le sens vertical et selon qu'elles sont munies ou non de repères avec ou sans chiffres [14, 25, 34].

## 2.2. EVALUATION QUALITATIVE

L'évaluation qualitative comporte deux étapes :

- un jugement d'experts
- un pré-test

Après rédaction, le questionnaire doit être relu par des experts du domaine. Ils apportent une appréciation subjective sur :

- . l'absence de dérive par rapport à l'objectif initial,
- . la planification du questionnaire ; il ne doit pas être une juxtaposition de questions mais un tout, les questions devant suivre un ordre logique,
- . la compréhension des questions par la population cible,
- . la non induction des réponses,
- . la pertinence des modalités de réponse,
- . le choix des réponses et leur nombre,
- . la non ambiguïté du vocabulaire,
- . l'utilité de toutes les questions afin de réduire au maximum la longueur du questionnaire [5, 17, 31].

Le pré-test est une étape à ne pas négliger. Il va permettre de mettre en lumière toutes les imperfections d'une première ébauche. Il sera une réelle mise à l'épreuve du questionnaire si il est pratiqué de manière rigoureuse. Il doit être administré à un échantillon représentatif de la population cible. Utiliser pour le pré-test des personnes disponibles, parce que hospitalisées ou participantes à une autre étude, fait courir le risque d'un biais de sélection [10].

Enfin, le pré-test doit être confié à un enquêteur éprouvé, doté de connaissances méthodologiques et scientifiques, capable de dépister les biais de réponses et de les expliquer [5].

### 2.3. TRADUCTION D'UN QUESTIONNAIRE

La traduction d'un questionnaire suppose deux étapes indispensables : une traduction littérale et une adaptation au contexte culturel, aux habitudes de vie et aux idiomes de la population cible [22, 38]. Cette seconde étape constitue la validation transculturelle du document. Elle est incontournable pour les items portant sur des dimensions sociales ou psychologiques [8, 21, 23, 27].

Une procédure de traduction des questionnaires s'est progressivement imposée dans le domaine bio-médical [22]. Elle est considérée par de nombreux chercheurs comme un outil permettant de limiter les problèmes de transfert culturel et comme une évaluation qualitative du questionnaire. Elle se déroule en quatre étapes.

Dans un premier temps, trois bilingues, de formation académique différente, traduisent le questionnaire indépendamment les uns des autres. Afin d'avoir une traduction de bonne qualité, les trois traducteurs ont, pour langue maternelle, la langue cible ; en outre, ils sont informés des objectifs de l'étude ainsi que des concepts sous-jacents aux items. La qualification et le niveau d'étude des traducteurs sont à considérer : un traducteur très cultivé risque d'utiliser un vocabulaire peu accessible à une population tout venant [22].

Dans une deuxième étape, les trois versions obtenues dans la langue cible sont contre-traduites dans la langue source par trois nouveaux bilingues. Ceux-ci possèdent la langue source comme langue maternelle. Contrairement à la première étape, il est ici préférable que les traducteurs ne connaissent ni les concepts sous-tendant le questionnaire, ni les objectifs de l'étude pour favoriser l'émergence des omissions et des erreurs de traduction.

La troisième étape est la rédaction d'une version provisoire du questionnaire. Pour cela, le comité de traduction, constitué des six traducteurs et/ou des experts du phénomène étudié et/ou des bilingues représentant la population cible, compare les trois contre-traductions à la version originale, identifiant ainsi les items dont les formulations divergent. Le comité de traduction a, alors, pour tâche d'élaborer, sur la base du consensus, une nouvelle rédaction dans la langue cible de ces items. Trois critères vont guider cette reformulation : une tournure grammaticale correcte, un vocabulaire accessible à tous et un style neutre.

La dernière étape ou pré-test, consiste à administrer le questionnaire à un groupe représentatif de la population cible. Après renseignement du questionnaire, l'enquêteur s'enquière des éventuelles difficultés à répondre. La rédaction définitive du questionnaire intègre le repérage de ces difficultés. Un autre pré-test consiste à administrer la version traduite ainsi que la version d'origine à un groupe de bilingues, afin qu'ils évaluent le degré de similitude entre les deux versions [22].

EUROQUEST, rédigé initialement en anglais, a été traduit en français. Traductions et contre-traductions ont été réalisées par deux linguistes, deux spécialistes en neurotoxicologie et deux

traducteurs culturellement proches des salariés auxquelles le questionnaire serait, par la suite, administré [8]. La version provisoire a été testée auprès d'un panel de 39 hommes et femmes représentant la population cible pour l'âge et la profession. Une réponse optionnelle, *Je ne sais pas*, a été temporairement incluse à chaque réponse. L'enquêtrice a ensuite recensé les difficultés rencontrées lors du renseignement du questionnaire, en particulier lorsque la réponse, *Je ne sais pas*, avait été choisie. Le comité de traduction a donc revu l'intitulé des questions mal comprises ou ambiguës et arrêté une version française définitive d'EUROQUEST, jointe en annexe [8].

A des fins de comparabilité des études entre elles, l'épidémiologiste utilise, autant que faire se peut, des questionnaires déjà employés dans d'autres études ; il doit s'assurer préalablement que la recherche des erreurs exposées dans le premier chapitre ainsi que le pré-test ont déjà été réalisés. Pour les questionnaires conçus dans une autre langue et traduits selon la méthodologie décrite plus haut, l'évaluation qualitative, même réalisée dans la langue source, suffit. Il reste cependant à réaliser, dans la langue cible, l'évaluation quantitative du questionnaire.

### 3. VALIDATION D'UN QUESTIONNAIRE

#### 3.1. "SCORING" : CONSTRUCTION D'UN SCORE

Un domaine peut exceptionnellement être exploré par un seul item ; en général, un domaine comprend plusieurs items ce qui nécessite de résumer l'information en construisant un score. Il existe plusieurs possibilités pour construire un score. La plus simple est d'effectuer la somme des notes attribuées aux différents items d'un même domaine [10]. L'autre consiste en une combinaison linéaire pondérée des items.

La pondération peut être décidée a priori par les experts. Quand les modalités de réponse aux items sont différentes à l'intérieur d'un même domaine, on peut standardiser les réponses afin de pouvoir calculer un score global par domaine, soit calculer plusieurs sous-scores dans un même domaine regroupant les items avec le même type de réponses [12, 36]. La pondération peut être décidée a posteriori après une analyse factorielle des correspondances (AFC).

EUROQUEST comporte 83 questions regroupées par les concepteurs en onze domaines (tableau 1). Le questionnaire offre trois types de réponses : la première mesure la fréquence des symptômes avec un choix de réponses entre, *Jamais ou très rarement, Quelquefois, Souvent, Très souvent*. Le second type recueille l'opinion des salariés par rapport aux symptômes décrits, le choix des réponses étant, *Pas du tout d'accord, Pas d'accord, D'accord, Tout à fait d'accord*. Le dernier type mesure la qualité de vie avec les possibilités telles que, *Très bon, Bon, Mauvais, Très mauvais*. L'analyse des 536 questionnaires montre une différence de répartition dans les réponses selon la modalité proposée (voir tableau 2). Pour le premier type, la réponse, *Jamais*, est majoritaire alors que, pour le deuxième type, les réponses se répartissent également sur les quatre options. Prenant en compte cette répartition des réponses, nous avons réduit, pour chaque item, les quatre choix à une alternative dichotomique, *Oui/Non*. Pour les items des domaines 1 à 8, *Non*, correspond à, *Jamais ou très rarement, Oui* à *Quelquefois/Souvent/Très souvent*. Pour les items des domaines 9 à 11, *Non*, correspond à, *Pas du tout d'accord/Pas d'accord*, et *Oui*, à, *D'accord/Tout à fait d'accord*. Pour le dernier type de réponses, la coupure s'est faite entre *Très bon* et les autres modalités de réponse. On a choisi ce regroupement de façon à équilibrer les réponses positives et négatives ce qui a pour effet d'augmenter la sensibilité.

Nous avons ensuite calculé comme score de chaque domaine, la somme des *Oui* à l'ensemble des questions du domaine considéré.

#### 3.2. EVALUATION QUANTITATIVE D'UN QUESTIONNAIRE

On appelle évaluation quantitative d'un questionnaire, l'exploration de ses qualités métrologiques : fiabilité, validité et sensibilité au changement [10, 12, 31]. COTTRAUX propose dans le tableau 3 une synthèse de tous les aspects métrologiques d'un questionnaire.

##### 3.2.1. ETUDE DE LA FIABILITE

Un questionnaire est fiable ou fidèle s'il donne des résultats comparables dans des situations comparables. Le concept de fiabilité correspond à celui de "précision" pour les scientifiques et est contenu dans celui de "dispersion" statistique. En épidémiologie, on parle encore de "reproductibilité de la mesure".

On distingue deux types de "fiabilité": la fiabilité liée à l'enquêteur et la fiabilité au cours du temps [10, 12, 35].

**La fiabilité liée à l'enquêteur** vérifie que les réponses de l'enquêté sont indépendantes de l'enquêteur ; elle s'étudie classiquement en calculant des coefficients de concordance inter-enquêteurs de Kappa ou des coefficients de corrélation intra-classes [15, 16, 36]. EUROQUEST étant un auto-questionnaire, cet aspect est sans objet.

**La fiabilité au cours du temps** contrôle la similitude des réponses d'un même individu lors de plusieurs passages du questionnaire, à des temps différents, mais dans des conditions expérimentales identiques. On étudie la corrélation statistique entre les scores obtenus lors des différentes passations du questionnaire [10, 18, 36]. Les tests de fiabilité d'EUROQUEST au cours du temps ont porté sur une population de 51 apprentis, 25 exposés et 26 "témoins", ayant répondu deux fois au questionnaire entre 1996 et 1997 ; ce laps de temps d'une part et le niveau de l'exposition d'autre part sont trop faibles pour envisager un quelconque effet neurotoxique des solvants. Entre 1996 et 1997, les scores sont restés stables : le pourcentage d'apprentis qui obtiennent le même score au même domaine est compris entre 49 et 76% selon le domaine considéré.

### 3.2.2. ETUDE DE LA VALIDITE

Un questionnaire est valide s'il mesure bien ce qu'il est censé mesurer [35].

Comme le montre une revue de la littérature portant sur l'évaluation de questionnaires médicaux, cet aspect est souvent négligé [10]. L'imprécision du concept même de validité rend compte en partie de cette négligence [17]. L'"American Psychological Association" a essayé d'en clarifier la définition en distinguant les validités de contenu, de structure et contre critère [1,9]. Mais le sens de ces trois termes n'est pas le même pour tous les auteurs et les divergences entre les différentes typologies de validation ne peuvent être entièrement expliquées par la diversité des formations académiques de leurs auteurs. Les tableaux 3, 4 et 5 synthétisent chaque point de vue.

Pour COTTRAUX, psychologue, la cohérence interne d'un questionnaire, recherchée notamment par le coefficient  $\alpha$  de CRONBACH, appartient à l'étude de la fiabilité (tableau 3) [12, 31] ; pour COSTE, statisticien, la cohérence interne est au contraire du domaine de la validité de contenu (tableau 4) [31]. Cet auteur différencie au sein de la validité de structure, une validité de structure interne, explorée notamment par les analyses en composantes principales et le coefficient  $\alpha$  de CRONBACH, et une validité de structure externe, explorée par la recherche de la validité convergente et divergente [10, 12]. Un troisième auteur, FERMANIAN, statisticien également, propose une typologie proche de celle de COSTE (tableau 5) [17]. Nous avons tenté une synthèse des trois sources précédentes, présentée dans le tableau 6 ; elle a essayé d'intégrer toutes les dimensions proposées par les trois auteurs ; elle nous a servi de référence pour étudier la validité d'EUROQUEST.

**La validité d'apparence** (Face Validity) résulte du jugement subjectif d'experts, ou simplement des utilisateurs tel qu'il est défini précédemment dans l'évaluation qualitative.

En ce qui concerne EUROQUEST, la validité apparente du questionnaire est garantie par l'évaluation qualitative que les concepteurs ont effectuée lors de son élaboration.

**La validité de contenu** (Content Validity) s'intéresse à la pertinence du questionnaire vis à vis du phénomène étudié. Cet aspect est à considérer si la définition du concept étudié ne fait pas l'objet d'un consensus. Elle permet de répondre aux questions suivantes :

- le phénomène exploré par le questionnaire correspond-il bien à ce que l'on recherche ?
- tous les domaines composant le phénomène étudié sont-ils bien pris en compte dans ce questionnaire ?

On peut rapprocher ces questions des termes de "pertinence" et "représentativité" utilisés par FERMANIAN pour définir respectivement la qualité des items et celle des domaines [17]. Pour cet auteur, un item est pertinent s'il correspond bien au domaine qu'il est censé explorer, ce qui permet de répondre à la première question. Un domaine est correctement représenté si il est exploré par un nombre d'items correspondant bien à son importance pour le phénomène étudié. Cela permet indirectement de répondre à la seconde question. En effet, la sur-représentation d'un domaine par un trop grand nombre d'items entraîne une sous-représentation des autres domaines, même si ceux-ci contribuent de façon prépondérante à la compréhension du phénomène étudié. Ainsi l'étude des troubles du sommeil est une composante importante du syndrome neurotoxique. Néanmoins le domaine "troubles du sommeil" ne doit pas être sur-représenté au détriment de celui des "symptômes neuro-sensoriels" par exemple. Nous n'avons pas étudié la validité de contenu d'EUROQUEST ; elle a, en effet, été évaluée dans la version anglaise et la traduction, du fait de la procédure suivie, permet de garantir la stabilité de contenu par delà le changement de langue.

**La validité contre critère** (Criterion Validity) évalue le degré de concordance entre les résultats fournis par le questionnaire et ceux fournis par un instrument de mesure du phénomène pris comme référence. Cet autre instrument doit être une référence indiscutable appelée "gold standard" dans la littérature anglo-saxonne ; il s'agit par exemple, de l'examen histo-pathologique dans le diagnostic de cancer.

Malheureusement il n'existe pas, pour tous les problèmes de santé, un outil diagnostique de certitude ; c'est le cas des maladies telles que la dépression, l'état anxieux, etc. Cependant certains questionnaires se sont imposés comme "gold standard" au fil des multiples études de validité dont ils ont fait l'objet. Ceci n'empêche pas qu'il soit nécessaire de construire un nouveau questionnaire, sur la dépression par exemple, car le "gold standard" s'avère, par exemple, trop long dans le contexte d'utilisation projeté.

La chronologie des mesures définit deux formes de validité contre critère :

- **validité concomitante** lorsque les mesures se font simultanément
- **validité prédictive** lorsque le critère de référence n'est pas immédiatement mesurable ; ainsi la question, *Avez-vous pleuré plusieurs fois par jour durant les trois derniers mois ?* de l'échelle GHQ (General Health Questionnaire) s'est révélé prédictive d'une tentative de suicide dans les six mois suivants [17, 31].

Faute d'outil de référence pour la mesure les effets neurotoxiques des solvants, la validité contre critère d'EUROQUEST n'a pu être testée. Nous nous sommes donc appliqués, comme cela est recommandé dans la littérature, à valider la "structure externe et interne" du questionnaire [9].

**La validité de structure** (Construct Validity) ; la structure "intime" d'un questionnaire peut être validée de façon indirecte :

- en utilisant des mesures du même phénomène, provenant d'autres sources que celles du questionnaire ; on parle alors de validité de structure externe,
- en utilisant les mesures même du questionnaire ; on parle alors de validité de structure interne.

***La validité de structure externe*** ; faute de "gold standard", on peut comparer les données du questionnaire avec celles provenant d'autres mesures et vérifier le degré de convergence ou de divergence des deux sources.

La validité convergente teste si les scores de différents domaines d'un questionnaire sont corrélés positivement à d'autres mesures des mêmes domaines. Elle a pu être réalisée pour deux domaines d'EUROQUEST : "mémoire" et "concentration", également mesurés par des tests neuropsychologiques ; ces tests font partie d'une batterie propre à la neurotoxicologie, NES (Neurobehavioural Evaluation System) [3]. La fonction mnésique a été explorée par des tests classiques : apprentissage d'associations de mots et rappel différé de ces associations (associate learning, recall), empan de chiffres dans l'ordre et rétrograde (digit span forwards et backwards) et test de reconnaissance de figures géométriques (pattern memory). La capacité de "concentration" a été évaluée par un test attentionnel très connu : le test de codage (digit symbol). Les salariés ont subi ces tests immédiatement avant de remplir EUROQUEST.

Nous avons testé l'hypothèse d'une corrélation négative entre les scores du questionnaire et les résultats de ces tests : plus les performances aux tests sont faibles, plus les salariés déclarent un nombre élevé de difficultés dans ce domaine ; le test statistique utilisé est le test de corrélation de PEARSON ; il n'existe pas de corrélation statistiquement significative ( $Rho = 0,17$ ) entre le domaine "concentration" et le test "digit symbol". Les corrélations entre le score "mémoire" du questionnaire et les performances aux différents tests de mémoire sont également négatives (exceptée celle avec la performance au "pattern memory") sans être statistiquement significatives. Seule la corrélation avec la performance au test "associate learning" est significative ( $Rho = - 0,21$ ,  $p = 0,02$ ). Le test "associate learning" est un test d'apprentissage de 2 mots associés, un prénom et un métier ; il explore la capacité à stocker une information en mémoire à moyen terme ce qui exige un effort conscient ; les autres tests mnésiques font appel à la mémoire à court terme qui est beaucoup plus "instinctive" et "inconsciente" ; la performance mnésique à court terme est souvent méconnue de la personne elle même. Ceci explique que les troubles de mémoire déclarés, donc auto-repérés, soient corrélés au seul test de mémoire "conscient" [28].

La validité divergente étudie l'absence de corrélation entre le score d'un domaine donné du questionnaire et une mesure qui, d'après les hypothèses d'études, ne peut avoir de lien théorique avec ce domaine [17, 18]. Pour EUROQUEST, on pourra, par exemple, tester l'hypothèse que les performances au test de vocabulaire, reflet du niveau socio-culturel, sont indépendantes des scores des domaines "troubles de l'humeur", ou "anxiété", ou "perception de la santé et de la vie".

La validité de structure externe d'un questionnaire ne peut s'évaluer sur une seule étude. Elle doit être confirmée ou infirmée au fil des études, menées par différents enquêteurs, testant chacun différentes hypothèses [17, 18]. Pour EUROQUEST, plusieurs hypothèses sont déjà envisagées, à titre d'exemple, citons la corrélation entre "fatigue" et test attentionnel, ou "troubles neurologiques" et vitesse de réaction, et celles citées ci-dessus dans le cadre de la validité divergente.

***La validité de la structure interne*** doit confirmer par les données, la structure postulée du questionnaire. Le lien existant entre deux items appartenant à un même domaine, donc explorant sur le plan théorique le même concept, doit être plus fort que le lien pouvant exister entre deux

items de domaines différents [17]. Cette validité s'explore classiquement par les analyses factorielles, exploratoires ou confirmatoires [10, 17].

Le coefficient  $\alpha$  de CRONBACH, fréquemment utilisé pour valider la structure interne (ou l'homogénéité interne) d'un questionnaire, est de plus en plus remis en cause dans la littérature [32]. Les critiques portent sur deux points. D'une part, la valeur du coefficient est liée au nombre d'items de l'échelle : à l'extrême le coefficient  $\alpha$  est maximal, égal à 1, lorsque le nombre d'items est très grand, quelles que soient les relations existantes entre eux. D'autre part, un coefficient  $\alpha$  élevé serait plus le reflet d'une certaine redondance parmi les items que la preuve d'une bonne cohérence interne de l'échelle [32].

Prenant en compte ces critiques, nous l'avons écarté au profit d'une AFC. Chacun des 11 domaines définis par les concepteurs et récapitulés dans le tableau 1 correspond-il bien à un axe de l'AFC ?

L'AFC n'a pas permis de retrouver la structure préalablement définie. Ainsi sur l'axe 1, on retrouve des questions appartenant à 7 domaines différents ; sur l'axe 10, 7 domaines sont représentés et sur l'axe 12, 9 domaines. Seuls les axes 2, 3, 4 et 11 semblent plus homogènes avec, sur chacun, des questions appartenant au maximum à 4 domaines différents. Cette AFC nous a cependant permis de faire l'hypothèse que les domaines d'EUROQUEST devaient être plus nombreux.

Nous les avons donc recherché au moyen de 2 autres méthodes statistiques descriptives, l'une est fondée sur l'odds-ratio entre 2 items, l'autre est une classification hiérarchique ascendante (CHA). La CHA a pour but de regrouper deux à deux les items dont le lien est plus fort qu'avec les autres items. S'agrègent ensuite, sur le même principe, un à un, les autres items. Les indices utilisés pour ces deux descriptions sont différents de la distance du  $\chi^2$  utilisé par l'AFC.

#### La méthode des odds-ratios

Elle teste l'association pour chaque couple d'items en calculant l'odds-ratio (OR). Pour des raisons de distribution, nous avons retenu le logarithme de l'OR,  $\log(\text{OR})$ , plutôt que l'OR lui-même. Ainsi une absence d'association se traduit par  $\log(\text{OR})$  nul, le signe du  $\log(\text{OR})$  est celui de l'association. On calcule ensuite les moyennes des  $\log(\text{OR})$  de chaque couple d'items d'un même domaine et appartenant à deux domaines différents.

Le tableau 7 récapitule toutes les possibilités de croisement :

- les moyennes des  $\log(\text{OR})$  intra-domaines, sur la diagonale
- les moyennes des  $\log(\text{OR})$  inter-domaines hors de la diagonale.

Nous retrouvons les moyennes les plus élevées sur la diagonale du tableau : les liaisons entre items sont donc en moyenne plus élevées au sein d'un même domaine qu'entre domaines.

Afin de mieux rendre compte des liaisons intra et inter-domaines, nous avons calculé, pour chaque domaine, la différence entre la moyenne des  $\log(\text{OR})$  intra-domaines et la moyenne des  $\log(\text{OR})$  inter-domaines. Une différence importante est le reflet d'une part, d'une faible liaison entre les questions du domaine et l'ensemble des questions des autres domaines, et d'autre part d'une forte liaison entre les items d'un même domaine ce qui traduit une bonne cohérence interne du domaine. Les différences s'échelonnent entre 0.13 et 0.89. Nous avons déterminé deux seuils et constitué trois groupes de domaines :

- différence  $\geq 0.8$  : forte cohérence interne du domaine
- $0.8 > \text{différence} \geq 0.5$  : bonne cohérence interne
- $0 < \text{différence} < 0.5$  : mauvaise cohérence interne du domaine comprenant des questions non spécifiques

Dans le groupe I, on retrouve 2 domaines : "anxiété" et "perception de la santé et de la vie".

Dans le groupe II, on retrouve 4 domaines : "troubles de l'humeur", "concentration", "mémoire", et "hypersensibilité".

Dans le groupe III, on retrouve 5 domaines : "fatigue", "symptômes aigus", "trouble du sommeil", "signes neurologiques" et "signes psychosomatiques".

Nous avons également tracé un graphique pour chaque question (tableau 8) où chaque point représente, pour chaque item, son  $\log(\text{OR})$  avec chacun des items. La simple lecture de ces graphiques devrait permettre de repérer visuellement les domaines, qui auraient dû être matérialisés par une ligne horizontale ; mais les graphiques sont au contraire en "dents de scie" témoignant d'une absence de proximité des items au sein d'un même domaine. Seul le groupe des questions "Hypersensibilité", cotées Q401 à Q408 sur le tableau 9, peut être individualisé parmi les 83 questions.

#### Classification hiérarchique ascendante et nouvelle conceptualisation

La deuxième méthode a combiné une nouvelle conceptualisation a priori du questionnaire et une CHA.

Nous avons fait l'hypothèse que la difficulté à retrouver la structure de départ en onze domaines d'EUROQUEST était liée à la conception même du questionnaire. Le découpage initial restitué de fait la teneur et le déroulement d'un interview clinique neurologique ; chaque domaine correspond à une fonction "théorique" du système nerveux (SN), basée sur l'activité produite, observable. Dans la réalité, chaque activité ne répond pas à une seule voie nerveuse ou à une zone cérébrale définie ; au contraire, une seule activité observable du SN met en jeu plusieurs structures ou voies cérébrales. Nous avons supposé l'existence d'une autre structure du questionnaire, qui, compte tenu du caractère exhaustif que les concepteurs ont voulu lui donner, pouvait être retrouvée tant sur le plan conceptuel que par les données.

Dans un premier temps, nous avons conçu un nouveau découpage du questionnaire. Pour ce faire, nous avons considéré les 2 listes, reconnues par la communauté scientifique, qui définissent les critères diagnostiques des manifestations neurotoxiques, syndrome psycho-organique et encéphalopathie toxique : la classification de Raleigh et la classification de l'OMS (Organisation Mondiale de la Santé) [7]. L'étude de ces classifications cliniques a permis de reconsidérer des domaines cliniques et de redistribuer les items en 14 domaines. Elle nous a amené à supprimer 7 questions qui ne correspondaient pas à des symptômes relevés dans les critères diagnostiques et/ou qui nous semblaient trop peu spécifiques, comme par exemple, *Au cours de ces derniers mois, avez-vous laissé tomber involontairement ce que vous aviez dans les mains ? Au cours de ces derniers mois, avez-vous eu les mains qui tremblaient ?*

Une analyse en CHA a ensuite été réalisée sur l'ensemble des 536 questionnaires, la distance étant définie par  $\sqrt{1 - r^2}$  où  $r$  est le coefficient de corrélation entre items et le critère d'agrégation retenu étant le critère de Ward ; elle a retrouvé 18 clusters à partir des 83 questions d'EUROQUEST (voir tableau 10). Les 14 domaines définis à partir des classifications cliniques sont bien retrouvés ; quatre domaines supplémentaires apparaissent (tableau 11). Par exemple, nous avons classé 6 questions se rapportant à la mémoire dans un domaine "mémoire" ; après la classification hiérarchique ascendante, nous l'avons subdivisé en deux sous-groupes : "troubles de mémoire subjectifs" et "troubles de mémoire objectivables". Les 7 items que nous avons, a priori, écartés se sont révélés être des questions n'appartenant à aucun des 18 clusters ou bien se liant tardivement à l'un de ces clusters.

La pertinence de ce nouveau découpage est confirmée par la reproductibilité des résultats ; en effet, la structure en 18 domaines est retrouvée dans les 2 sous-populations de l'étude, "Apprentis" et "Toluène".

Nous avons ensuite repris, en considérant les 18 domaines, la première démarche d'analyse décrite précédemment avec le calcul des moyennes des logOR, étudiant les liaisons entre items inter et intra-domaines. Les moyennes les plus élevées se trouvent toujours sur la diagonale du tableau, comme avec l'ancienne classification (voir tableau 12). Si l'on utilise, pour classer les domaines selon leur cohérence interne, les mêmes seuils que lors de la première analyse, on retrouve :

- dans le premier groupe de forte cohérence, 5 domaines sur 18 soit 28% des domaines (contre 2 sur 11 soit 18% lors de la première analyse),
- dans le groupe intermédiaire : 11 des 18 domaines, soit 61% (contre 4 sur 11 soit 36%)
- dans le groupe de faible cohérence , 2 sur 18 soit 11% contre 5 sur 11 soit 45% lors de la première analyse.

Cette nouvelle définition des domaines a donc amélioré la cohérence interne du questionnaire.

### 3.2.3. ETUDE DE LA SENSIBILITE AU CHANGEMENT

Un questionnaire est sensible lorsqu'il est capable de mesurer des différences minimales entre individus ou chez le même individu au cours du temps [10, 17, 18, 31]. Ainsi une échelle destinée à mesurer la mobilité de l'articulation de la hanche chez des malades souffrant d'arthrose doit pouvoir refléter l'apparition de raideurs comme l'amélioration de la mobilité. Plus le questionnaire est sensible mieux il met en évidence des changements, même minimes, de cette mobilité.

Les données que nous possédons ne nous permettent pas de rechercher pour le moment la sensibilité au changement d'EUROQUEST, l'effet neurotoxique des solvants n'ayant pas été mis en évidence ; nous étudierons cet aspect dans une population "atteinte".

## 4. DISCUSSION

La discussion s'articule autour des 4 points suivants :

- La confrontation des disciplines psychosociologiques et biomédicales est-elle féconde ?
- Quels enseignements tire-t-on de la mise à l'épreuve des méthodes d'évaluation sur EUROQUEST ?
  - . Pourquoi la structure originelle d'EUROQUEST n'a pu être validée ?
  - . Intérêt et limites des méthodes : odds-ratios et classification hiérarchique ascendante
- Quelle stratégie adopter pour une évaluation métrologique ?
- Quelles sont les conséquences de la première évaluation d'EUROQUEST ?

### **La confrontation des disciplines psychosociologiques et biomédicales est-elle féconde ?**

L'étude du même champ dans deux disciplines différentes est intéressante.

La psychosociologie est riche d'une tradition éprouvée dans l'élaboration, sachant soignée la formulation des questions ; les questionnaires psychosociologiques explorent un même domaine avec de multiples items afin d'en augmenter la cohérence interne. En outre, les phénomènes étudiés, plus complexes par nature, ne peuvent se réduire à une dichotomie et sont mesurés au moyen d'échelles quantitatives.

A l'inverse, la recherche biomédicale très imprégnée de la tradition clinique, réduit souvent le phénomène de santé à une appréciation de type diagnostique "malade/non malade", qui est très opératoire pour une décision thérapeutique mais réductrice en situation de recherche. Cependant l'évolution tend à substituer cette approche dichotomique au profit d'une quantitation des phénomènes de santé mesurés ; ce qui est non seulement plus adapté au contexte de la recherche mais également amplement justifié par l'évolution réelle de la plupart des phénomènes de santé : un continuum progressif de la bonne santé à la maladie. Quant à la formulation des questions, le questionnaire médical est moins soucieux de respecter la façon dont les personnes expriment leurs symptômes ; il dérive d'une tradition clinique qui est "l'interrogatoire" médical où le médecin dirige l'entretien. En ce qui concerne le nombre d'items nécessaire à l'exploration d'un domaine, la tendance est plutôt à la limitation maximale afin de faciliter la passation du questionnaire [11, 39]. A l'extrême, chaque domaine n'est exploré que par un seul item. Ce type d'échelles est d'un usage facile car elles sont rapidement renseignées mais elles manquent de fiabilité car les mesures sont entachées d'une grande variance et le pouvoir discriminant de telles échelles est donc faible [39].

Au crédit des disciplines biomédicales, rappelons l'apport d'une méthodologie de traduction des questionnaires rigoureuse, intéressante à considérer quand on veut comparer les différentes versions linguistiques d'un même questionnaire.

La méthodologie de validation quantitative, originaire des sciences sociales, est reprise depuis quelques années par la recherche biomédicale. C'est le domaine où les deux disciplines se rejoignent. Cependant leur utilisation plus généralisée se heurte à l'absence de consensus sur les concepts de validité et à l'absence de normalisation des méthodes statistiques.

## **Pourquoi la structure originelle d'EUROQUEST n'a pu être validée ?**

Une exposition prolongée à des substances neurotoxiques est responsable de lésions touchant à la fois le système nerveux central et périphérique [26]. Les symptômes qui en découlent sont très variés. Les classifications du syndrome neurotoxique retrouvées dans la littérature, classification de l'OMS (1985), classification de Raleigh (1986) montrent en effet que les symptômes n'apparaissent pas système par système. Dès le début de l'atteinte, les personnes peuvent présenter des lésions motrices, sensibles et/ou cognitives, dont la gravité dépendra de la durée ainsi que de l'intensité de l'exposition [7].

La structure originelle d'EUROQUEST regroupe les items par sous-système : système neurologique périphérique (questions portant sur des atteintes motrices et/ou sensibles), système sympathique, sensoriel, atteintes cognitives (troubles de concentration, troubles de mémoire...), troubles de la personnalité, etc.

Le nouveau découpage d'EUROQUEST a utilisé un autre critère ; les items ont été regroupés non pas sur la base fonctionnelle mais sur une base anatomo-pathologique. Il a été confirmé par l'analyse en classification hiérarchique ascendante.

Nous avons finalisé le découpage en domaines en mettant en relation les domaines de la CHA à ceux décrits par les classifications de Raleigh et de l'OMS.

Même avec un nouveau découpage, l'AFC n'a pu confirmer la nouvelle structure du questionnaire.

L'AFC regroupe les items par axe, chaque axe étant une combinaison linéaire de tous les items du questionnaire. L'axe principal est celui qui permet d'expliquer le pourcentage maximal de variance des réponses au sein de la population étudiée. Ceci pourrait peut-être permettre de comprendre pourquoi la nouvelle structure d'EUROQUEST, de même que la structure originelle, n'a pas pu être validée par une AFC. La majorité des salariés de l'étude n'était exposée, que depuis peu de temps, aux neurotoxiques et donc dépourvue, pour la majorité d'entre eux, d'atteintes du SNC [26]. La population d'étude était donc très proche d'une population "normale". Dans cette situation, la variance des réponses est plus faible que dans une population comportant des personnes saines, peu et fortement atteintes.

Il semble donc que la faible variance des réponses rend l'AFC inadaptée à la validation de la structure interne d'un questionnaire.

Kiesswetter, après avoir évalué une nouvelle version du « Psychological-Neurological Questionnaire » [36], questionnaire ayant les mêmes objectifs qu'EUROQUEST, conclue à la nécessité de développer un outil de dépistage des syndromes neurotoxique plus sensible [6, 24].

## **Intérêt et limites des méthodes : odds-ratios et classification hiérarchique ascendante**

Par contre, les nouvelles méthodes statistiques appliquées à cette occasion ont montré leur intérêt. Le calcul des OR entre items a permis un "diagnostic" de cohérence interne pour chaque domaine ; ainsi peut-on identifier les domaines très, bien ou peu "cohérents". Il s'agit d'une méthode descriptive simple, permettant une représentation graphique, visuellement interprétable, à mettre en œuvre en première intention. Cette méthode n'a d'autres hypothèses que celles de la structure postulée du questionnaire. L'analyse en CHA permet par contre de concevoir, à partir des données, un nouveau découpage d'un questionnaire en domaines. Sa fragilité tient à l'instabilité des données. Ce découpage ne peut être validé que par la confrontation à un modèle théorique et n'être confirmé que sur plusieurs jeux de données.

## Quelle stratégie adopter pour une évaluation métrologique ?

La validation d'un questionnaire devrait apparaître en épidémiologie comme une étape aussi indispensable que l'étalonnage et la validation d'un nouvel instrument de mesure dans les sciences physiques.

Les évaluations de questionnaire publiées ces dernières années dans la littérature biomédicale ont montré des lacunes majeures dans ce domaine ; les techniques statistiques ne sont pas employées à bon escient, des qualités primordiales d'un questionnaire ne sont pas explorées ou étudiées de façon appropriée [10].

La validation d'un questionnaire est un travail important qui demande l'emploi de techniques statistiques précises. En théorie, elle doit balayer tous les aspects métrologiques du questionnaire : fiabilité, validité et sensibilité au changement.

En pratique, l'étude complète de ces trois rubriques est rarement faisable par manque de temps et/ou de données. Faut-il alors privilégier l'étude de l'une de ces rubriques au dépend des autres ? Actuellement, les chercheurs qui vont utiliser un nouveau questionnaire portent essentiellement leurs efforts sur l'étude de la fiabilité [10], plus facile à explorer que la validité [36]. Malheureusement, la démonstration d'une grande fiabilité ne peut pas compenser une faible validité. On mesure, avec un tel questionnaire, un concept qui n'est pas le concept cible mais avec beaucoup de reproductibilité. Une mesure valide à fiabilité réduite est plus utile qu'une mesure très fiable d'un concept éloigné de l'objet de l'étude.

Par contre, on peut opposer fiabilité au cours du temps, et sensibilité au changement [12]. On étudiera l'une ou l'autre de ces qualités selon le type de questionnaire employé. Si l'échelle est supposée mesurer un trait permanent du caractère d'un individu, comme la "recherche de risque", on privilégiera la fiabilité au cours du temps. Si par contre, le phénomène mesuré varie au cours du temps, comme par exemple l'état dépressif, on étudiera la sensibilité au changement du questionnaire [12, 39]. Cette qualité est primordiale lorsque le questionnaire est employé comme outil de dépistage [28]

De même, on ne recherchera pas le même type de validité lorsque le questionnaire explore un seul domaine, par exemple la mobilité articulaire chez des arthrosiques, et lorsqu'il explore plusieurs domaines : la mobilité, la douleur et la gêne fonctionnelle. Dans le premier cas, la validité de structure, et plus particulièrement l'homogénéité (ou cohérence interne), sera la qualité la plus importante du questionnaire : chaque item doit bien explorer la mobilité des articulations et non la gêne fonctionnelle [39]. Pour le second cas, la cohérence interne jouera un rôle moindre ; la validité de contenu, pertinence et représentativité des items, sera par contre à étudier avec attention.

## Quelles sont les conséquences de la première évaluation d'EUROQUEST ?

La validation d'EUROQUEST a concerné la fiabilité, une partie de la validité, la validité de structure, externe pour deux domaines, et interne.

La recherche pour un questionnaire d'une cohérence interne plus forte que celle postulée à la conception était faisable dans le cas d'EUROQUEST, car les concepteurs ont voulu un questionnaire exhaustif sur le plan des fonctions du système nerveux explorées. Mais la recherche d'une cohérence interne plus forte ne peut s'appuyer uniquement sur des méthodes statistiques. En effet une forte corrélation entre items peut n'avoir aucune explication biologique ou physiologique. Il est recommandé de confronter les résultats des tests statistiques aux

connaissances cliniques. Le chercheur peut ainsi définir des domaines fondés sur des hypothèses biologiques plausibles et dont la cohérence interne est acceptable [10, 39].

Cette première évaluation impliquerait une nouvelle conception du questionnaire avec la suppression de 7 questions peu informatives et l'utilisation du nouveau découpage en domaines. Cependant, avant de modifier le questionnaire, il s'agira de compléter l'évaluation en particulier, dès que les données seront disponibles, par la mesure de la sensibilité au changement. A la lumière d'une évaluation complète, nous pourrions ensuite envisager la rédaction de la deuxième version d'EUROQUEST.

## 5. CONCLUSION

L'élaboration et la validation des questionnaires dans les sciences biomédicales peuvent profiter des acquis des sciences sociales. Cependant des adaptations sont nécessaires. En particulier les méthodes de validation classiques peuvent se révéler inadéquates. Tel a été le cas de la validation initiale d'un questionnaire de symptômes neurotoxiques, EUROQUEST. Devant cette situation, il faut savoir recourir à d'autres méthodes propres à chaque situation et de ce fait difficilement normalisables.

## BIBLIOGRAPHIE

1. American Psychological Association. Standards for educational and psychological testing. American Psychological Association, Washington DC, 1985.
2. Bacher, F. Les enquêtes en psychologie - Première description. Lille : Presse Universitaire de Lille ; 1982.
3. Baker, E.L. ; Letz, Re. ; Fidler, A.T. A computer-based neurobehavioural evaluation system for occupational and environmental epidemiology: methodology and validation studies. Neurobehav. Toxicol. Teratol. 7 : 369-377 ; 1985.
4. Blanchet, A. ; Ghiglione, R. ; Massonnat, J. ; Trognon, A. Les techniques d'enquête en sciences sociales. Paris : Dunod ; 1987 : 191.
5. Baudier, F. Education pour la Santé. Guide pratique. Besançon : SDES ADEES ; 1987 : 278.
6. Bukowski, JA. ; Sargent, EV. ; Pena, BM. Evaluation of the utility of standard history questionnaire in assessing the neurological effects of solvents. Am. J. Ind. Med. 22 (3) : 337-345 ; 1992.
7. Chang, L.W. ; Dyer, R.S. Handbook of Neurotoxicology, New-York : 340 ; 1995.
8. Chouaniere, D. ; Cassitto, M.G. ; Spurgeon, A. ; Verdier, A. ; Gilioli, R. An International Questionnaire to Explore Neurotoxic Symptoms. Environ. Res. 73 : 70-72 ; 1997.
9. Colvez A., Gardent H, Bucquet D. Validation d'indicateurs en gérontologie : application au domaine des indicateurs de capacité fonctionnelle. Les indicateurs d'incapacité fonctionnelle en Gérontologie. Information, validation, utilisation. INSERM : Réseau INSERM de Santé Publique : 37-77.
10. Coste, J. ; Fermanian, J. ; Venot, A. Methodological and statistical problems in the construction of composite measurement scales: a survey of six medical and epidemiological journals. Stat. Med. 14 : 331-345 ; 1995.
11. Coste, J. ; Guillemin, F. ; Pouchot, J. ; Fermanian, J. Methodological Approaches to Shortening Composite Measurement scales. J. Clin. Epidemiol. 50 (3) : 247-252 ; 1997.
12. Cottraux, J. ; Bouvard, M. ; Legeron, P. Méthodes et Echelles d'évaluation des comportements. Paris : EAP : 22-27.
13. De Ketele, J.M. ; Roegiers, X. Méthodologie du recueil d'informations. Fondements des méthodes d'observations, de questionnaires, d'interviews et d'études de documents. Bruxelles : DeBoeck ; 1993 : 226.

14. Dixon, JS. ; Bird, HA. Reproducibility along a 10 cm vertical visual analogue scale. *Ann. Rheum. Dis.* 40 : 87-89 ; 1981.
15. Fermanian, J. Mesure de l'accord entre deux juges : cas qualitatif. *Rev. Epidém. et Santé Publ.* 32 : 140-147 ; 1984.
16. Fermanian, J. Mesure de l'accord entre deux juges : cas quantitatif. *Rev. Epidém. et Santé Publ.* 32 : 408-413 ; 1984.
17. Fermanian, J. Evaluer correctement la validité d'une échelle : les nombreux pièges à éviter. *Rev. Epidém. et Santé Publ.* 44 : 278-286 ; 1996.
18. Fermanian, J. Methodology and statistical problems in the assesment of memory by rating scales. *Arch. Gerontol. Geriatr. Suppl.* 1 : 215-221 ; 1989.
19. Ghiglione, R. ; Matalon, B. Les enquêtes sociologiques, théorie et pratique. Paris : Armand Colin ; 1970 : 297.
20. Giacobbi, M. ; Roux, J.P. Initiation à la sociologie. Les grands thèmes, la méthode, les grands sociologues. Paris : Hatier ; 199 : 187-194.
21. Gilioli, R. EURONEST: A concerted action of the European Community for the study of organic solvents neurotoxicity. *Environ. Res.* 69 : 89-98 ; 1993.
22. Guillemin, F. ; Bombardier, C. ; Beaton, D. Cross-cultural adaptation of health-related quality of life mesures: literature review and propose guidelines. *J. Clin. Epidemiol.* 46 : 1417-1432 ; 1993.
23. Guillemin, F. ; Paul-Dauphin, A. ; Virion, JM. ; Bouchet, C. ; Briançon, S. Profil de santé DUKE : adaptation culturelle et validation d'une mesure de qualité de vie générique. *Rev. Epidém. et Santé publ.* 43, Résumé des communications ADELFF, Nancy 1995, CO4, 1995.
24. Iregren, A. Subjective and objective signs of organic solvent toxicity among occupationnally exposed workers. An experimental evaluation. *Scand. J. Work Environ. Health* 12 (5) : 469-475 ; 1986.
25. Joos, E. ; Peretz, A. ; Beguin, S. ; et al. Reliability and reproductibility of analogue scale and numeric rating scale for therapeutic evaluation of pain in rheumatic patient. *J. Rheumatol.* 18 : 1269-1270 ; 1991.
26. Kiesswetter, E. ; Sietmann, B. ; Seeber, A. Standardization of a questionnaire for neurotoxic symptoms. *Environ. Res.* 73 : 73-80 ; 1997.
27. Leplege, A. ; Mesbah, M. ; Ecosse, E. Evaluation des propriétés psychométriques de la version française d'un questionnaire de mesure de la qualité de vie : le MOS SF-36 (version 1.3). *Rev. Epidém. et Santé Publ.* 43, Résumé des communications ADELFF, Nancy 1995, CO3, 1995.

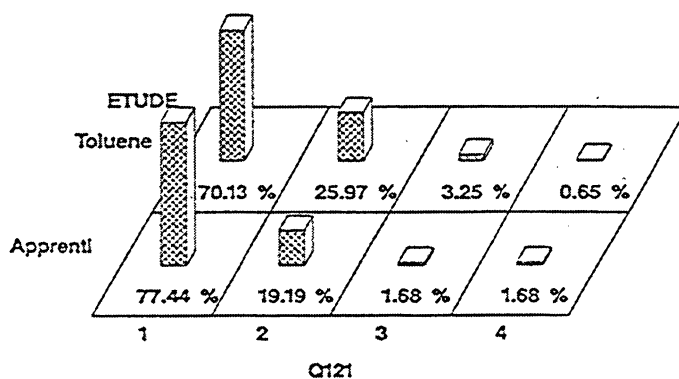
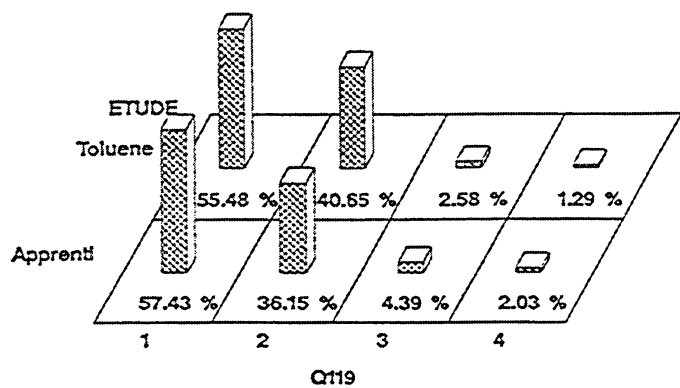
28. Lundberg, I. ; Högberg, M. ; Michélsen, H. ; Nise, G. ; Hogstedt, C. Evaluation of the Q16 questionnaire on neurotoxic symptoms and review of its use. *Occup. Environ. Med.* 54 : 343-350 ; 1997.
29. Massin, N.; Bohadana, B.; Wild, P.; Goutet, P.; Kirstetter, H.; Toamain, JP. Airway responsiveness, respiratory symptoms, and exposures to soluble oil mist in mechanical workers. *Occup. Environ. Med.* 53 : 748-752 ; 1996.
30. McHorney, C. ; Ware, J. ; Raczek, A. The MOS 36-Item Short-Form health survey (SF-36): II. Psychometric and clinical test of validity in measuring physical and mental health constructs. *Medical Care.* 31 (3) : 247-263 ; 1993.
31. Moret, L. ; Chwalow, J. ; Baudouin-Balleur, C. Evaluer la qualité de vie : construction d'une échelle. *Rev. Epidém. et Santé Publ.* 41 : 65-71 ; 1993.
32. Moret, L. ; Mesbah, M. ; Chwalow, J. ; Lellouche, J. Validation interne d'une échelle de mesure. Relation entre analyse en composante principale, coefficient  $\alpha$  de Cronbach et coefficient de corrélation intraclasse. *Rev. Epidém. et Santé Publ.* 41 : 179-186 ; 1993.
33. Muchelli, R. Le questionnaire dans l'enquête psychosociale. Paris : ESF (Formation permanente en sciences humaines) 9<sup>ème</sup> ed ; 1989.
34. Paul-Dauphin, A. ; Guillemin, F. ; Briançon, S. Le choix des échelles visuelles analogiques dans les études longitudinales. *Rev. Epidém. et Santé Publ.* 43, Résumé des communications ADELFF, Nancy 1995, Ca5, 1995.
35. SANTE PUBLIQUE. Extrait : "Le questionnaire Médical". Paris : INSERM ; 1978 : 11-21.
36. Seeber, A. ; Blaszkewicz, M. ; Golka, K. ; Kiesswetter, E. Solvent Exposure and rating of well-being: dose-effect relationships and consistency of data. *Environ. Res.* 73 : 81-91 ; 1997.
37. Sellitz, C. ; Wrightsman, I.S. ; Cook, S.W. Les méthodes de recherche en sciences sociales. Edition HRW.
38. Verdier, AS. La traduction de questionnaires de qualité de vie : spécificité de la tâche et cahier des charges. *Journal d'économie médicale, hors série. Résumé des communications, Santé et qualité de vie : concepts, mesures et politiques*, Paris, 69, 1996.
39. Wright, JG. ; Feinstein, AR. A comparative contrast of clinimetric and psychometric methods for constructing indexes and rating scales. *J. Clin. Epidemiol.* 45 (11) : 1201-1218 ; 1992.

**TABLEAUX 1 à 12**

## Tableau 1. Domaines d'origine d'EUROQUEST

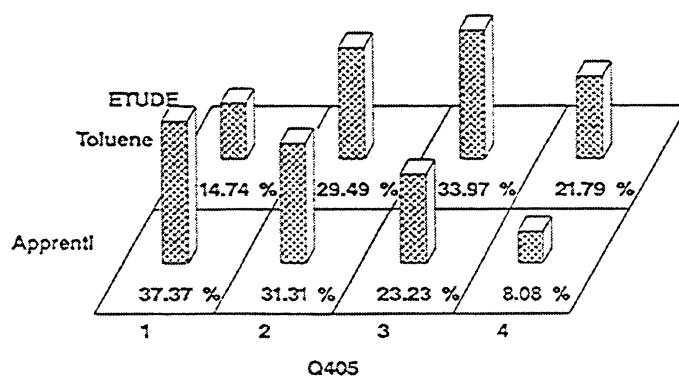
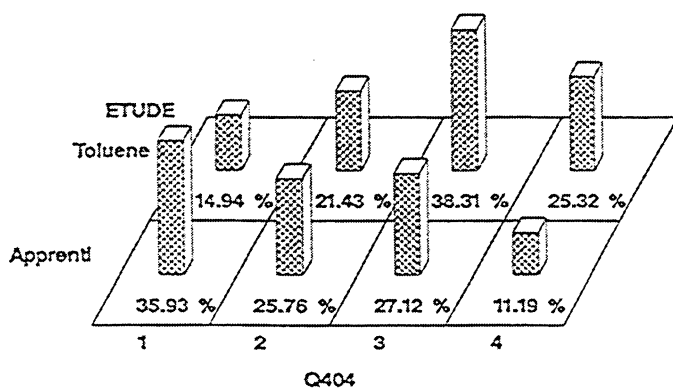
- 1 - *symptômes neurologiques (11 items)*
- 2 - *symptômes psychosomatiques (15 items)*
- 3 - *troubles de l'humeur (11 items)*
- 4 - *troubles de mémoire (6 items)*
- 5 - *troubles de la concentration (4 items)*
- 6 - *fatigue (7 items)*
- 7 - *troubles du sommeil (5 items)*
- 8 - *symptômes aigus (6 items)*
- 9 - *hypersensibilité à l'environnement (8 items)*
- 10 - *anxiété (perception de la santé) (6 items)*
- 11 - *perception de la santé et de la qualité de vie (4 items)*

Tableau 2. Répartition des réponses selon la modalité de réponse proposée



Modalités de fréquence :

- 1 Jamais ou très rarement
- 2 Quelquefois
- 3 Souvent
- 4 Très souvent



Modalités de fréquence :

- 1 Pas du tout d'accord
- 2 Pas d'accord
- 3 D'accord
- 4 Tout à fait d'accord

**Tableau 3. Validation quantitative d'un questionnaire selon Cottraux [12]**

Validité <i>Validity</i>	Fiabilité (ou fidélité) <i>Reliability</i>	Sensibilité au changement <i>Sensitivity</i>
<p><b>1 - Validité de contenu</b> (ou pragmatique) <i>Content / face validity</i> Adéquation aux situations fréquemment rencontrées par le sujet Adéquation et exhaustivité avec laquelle un domaine spécifique est couvert</p> <p><b>2 - Validité sur critère</b> <i>Criterion validity</i>: Concordance entre les résultats et les observations mesurant le phénomène avec exactitude</p> <ul style="list-style-type: none"> <li>- empirique : Correspondance des résultats aux classifications cliniques</li> <li>- concomitante : Corrélation de l'instrument à d'autres formes de mesures des mêmes variables</li> <li>- concurrente : Corrélation de l'instrument à d'autres instruments du même type déjà validés</li> <li>- prédictive : Possibilité d'émettre un pronostic à partir des résultats ; la mesure proposée prévoit une future valeur critère</li> </ul> <p><b>3 - Validité de structure</b> <i>Construct validity</i> Possibilité de mettre en correspondance les résultats avec une théorie préalable</p>	<p>Stabilité avec laquelle un questionnaire produit une même réponse à une question posée dans une même situation</p> <p><b>1 - Fiabilité de l'enquêteur</b> (ou interjuges) : Corrélation obtenue par deux (ou plusieurs) observateurs avec le même instrument</p> <p><b>2 - Fiabilité au cours du temps</b> (test-retest) : Corrélation des résultats de deux (ou plusieurs) évaluations dans les mêmes conditions expérimentales mais à des temps différents</p> <p><b>3 - Fiabilité de l'instrument</b> (cohérence interne, homogénéité) corrélation des différents items entre eux ainsi que corrélation avec une note totale</p>	<p><b>1 - Sensibilité interindividuelle</b> Capacité de l'instrument à discriminer des individus différents</p> <p><b>2 - Sensibilité intraindividuelle</b> Capacité de l'instrument à détecter des différences chez un même sujet</p>

**Tableau 4. Etude de la validité d'un questionnaire d'après Coste [10]**

Validité de contenu <i>Content validity</i>	Validité contre critère <i>Criterion validity</i>	Validité du construit/de structure <i>Construct validity</i>
Adéquation, exhaustivité avec lesquels le domaine étudié est décrit étude de la littérature, étude de cas cliniques	Validité contre critère - predictive - concurrente	<p><b>Interne :</b></p> <p><i>Factorial validity</i> (étudiée grâce à des analyses descriptives multivariées)</p> <p><i>Internal consistency</i> (cohérence interne)</p> <p><b>Externe :</b></p> <p>- <i>convergent validity</i> : corrélation avec les résultats d'une autre méthode mesurant le même phénomène</p> <p>- <i>discriminant validity</i> : absence de corrélation avec les résultats d'une autre méthode connue pour être sans relation avec la structure évaluée</p> <p>Mesure par rapport à une ou plusieurs variables explicatives</p> <p>- <i>Incremental validity</i> : méthode de mesure supérieure à toute mesure préexistante</p>

**Tableau 5. Etude de la validité d'un questionnaire selon Fermanian [17]**

<b>Validité d'apparence</b> <i>Face validity</i>	<b>Validité de contenu</b> <i>Content validity</i>	<b>Validité contre critère</b> <i>Criterion validity</i>	<b>Validité du construit</b> <i>Construct validity</i>
Jugement d'expert sur la valeur apparente de l'échelle	<p><b>Pertinence :</b></p> <p>chaque item appartient bien à un domaine étudié</p> <ul style="list-style-type: none"> <li>- Avis d'expert ; kappa, kappa pondéré, coefficient de corrélation intraclassé entre experts</li> <li>- Banque d'items, théorie des facettes de Guttman lorsque le concept est "flou"</li> </ul> <p><b>Représentativité :</b></p> <ul style="list-style-type: none"> <li>- Représentation de chaque domaine fonction de son importance dans l'"univers"</li> <li>- Jugement d'expert</li> </ul>	<p><b>Validité concomitante</b></p> <p><i>concurrent validity</i> : échelle et critère de référence utilisés en même temps</p> <ul style="list-style-type: none"> <li>- <b>Validité prédictive</b></li> </ul> <p><i>predictive validity</i> : phénomène est mesuré d'abord puis après un certain temps, le critère de référence</p>	<p>Test des hypothèses basées sur la conception théorique du phénomène mesuré par l'échelle</p> <p>(Analyses factorielles : Analyse des correspondances Analyses en composantes principales)</p> <p><b>Validité convergente</b></p> <p><b>Validité divergente</b> (coefficient de corrélation de Pearson ou non paramétrique)</p>

**Tableau 6. Proposition pour étudier la validité d'un questionnaire :**  
**synthèse des tableaux 3, 4 et 5**

<b>Validité d'apparence</b> <i>Face validity</i>	<b>Validité de contenu</b> <i>Content validity</i>	<b>Validité contre critère</b> <i>Criterion validity</i>	<b>Validité du construit</b> <i>Construct validity</i>
Jugement d'expert sur la valeur apparente de l'échelle	<p><b>Pertinence :</b>  chaque item appartient bien à un domaine étudié</p> <ul style="list-style-type: none"> <li>- Avis d'expert ; kappa, kappa pondéré, coeff. de corrélation intraclasse entre experts</li> <li>- Banque d'items, théorie des facettes de Guttman lorsque le concept est "flou"</li> </ul> <p><b>Représentativité :</b></p> <ul style="list-style-type: none"> <li>- Représentation de chaque domaine en fonction de son importance dans l'"univers"</li> <li>- Jugement d'expert</li> </ul>	<p><b>Validité concomitante :</b>  <i>concurrent validity</i> : échelle et critère de référence utilisés en même temps</p> <p><b>Validité prédictive :</b>  <i>predictive validity</i> : le phénomène est mesuré d'abord puis après un certain temps le critère de référence</p>	<p><b>Interne :</b></p> <ul style="list-style-type: none"> <li>- <i>Factorial validity</i> (Analyses descriptives multivariées)</li> <li>- <i>Internal consistency</i> : cohérence interne</li> </ul> <p><b>Externe :</b></p> <ul style="list-style-type: none"> <li>- <i>convergent validity</i> : corrélation avec les résultats d'une autre méthode mesurant le même phénomène</li> <li>- <i>discriminant validity</i> : absence de corrélation avec les résultats d'une autre méthode connue pour être sans relation avec la structure évaluée</li> <li>Mesure par rapport à une ou plusieurs variables explicatives</li> </ul>

Tableau 7. Moyenne des log(OR) entre deux items appartenant au même domaine ou à l'un des dix autres domaines

D1	D2												
	S.Neu- rolog- iques	S.Psy- choso- matiq- ues	Troub- les Humeur	Troub- les Memoi- re	Troub- les Conce- ntration	Troub- les Fatig- ue	Troub- les Somme- il	Sympt- omes Aigus	Hyper- sensi- bilite	Anxie- te V	Anxie- te VI	ALL	
	LOG_OR	LOG_OR	LOG_OR	LOG_OR	LOG_OR	LOG_OR	LOG_OR	LOG_OR	LOG_OR	LOG_OR	LOG_OR	LOG_OR	LOG_OR
	MEAN	MEAN	MEAN	MEAN	MEAN	MEAN	MEAN	MEAN	MEAN	MEAN	MEAN	MEAN	MEAN
S.Neurologiques	1.073	0.826	0.838	0.913	0.999	0.776	0.423	0.757	0.390	0.680	0.782	0.783	
S.Psychosomatiques	0.826	0.909	0.879	0.816	0.935	0.850	0.496	0.845	0.475	0.627	0.594	0.775	
Troubles Humeur	0.838	0.879	1.461	0.881	1.350	0.974	0.506	0.788	0.618	0.788	0.772	0.909	
Troubles Memoire	0.913	0.816	0.881	1.471	1.173	0.869	0.532	0.696	0.562	0.705	0.681	0.834	
Troubles Concentration	0.999	0.935	1.350	1.173	1.752	1.064	0.581	0.887	0.526	0.893	0.966	0.991	
Fatigue	0.776	0.850	0.974	0.869	1.054	1.308	0.429	0.835	0.455	0.689	0.778	0.821	
Troubles Sommeil	0.423	0.496	0.506	0.532	0.581	0.429	0.736	0.381	0.410	0.535	0.367	0.480	
Symptomes Aigus	0.757	0.845	0.788	0.696	0.867	0.835	0.381	1.085	0.279	0.481	0.556	0.706	
Hyper.sensibilite	0.390	0.475	0.618	0.562	0.526	0.455	0.410	0.279	1.290	0.420	0.407	0.534	
Anxiete V	0.680	0.627	0.788	0.705	0.893	0.689	0.535	0.481	0.420	1.582	0.683	0.704	
Anxiete VI	0.782	0.594	0.772	0.681	0.966	0.778	0.367	0.556	0.407	0.683	1.577	0.691	
ALL	0.783	0.775	0.909	0.834	0.991	0.821	0.480	0.706	0.534	0.704	0.691	0.757	
Difference between the global score and the domain score	0.29	0.13	0.55	0.64	0.76	0.49	0.26	0.38	0.76	0.88	0.99		

Tableau 8. Graphique des log(OR)  
entre les items Q101 à Q109 et les autres items

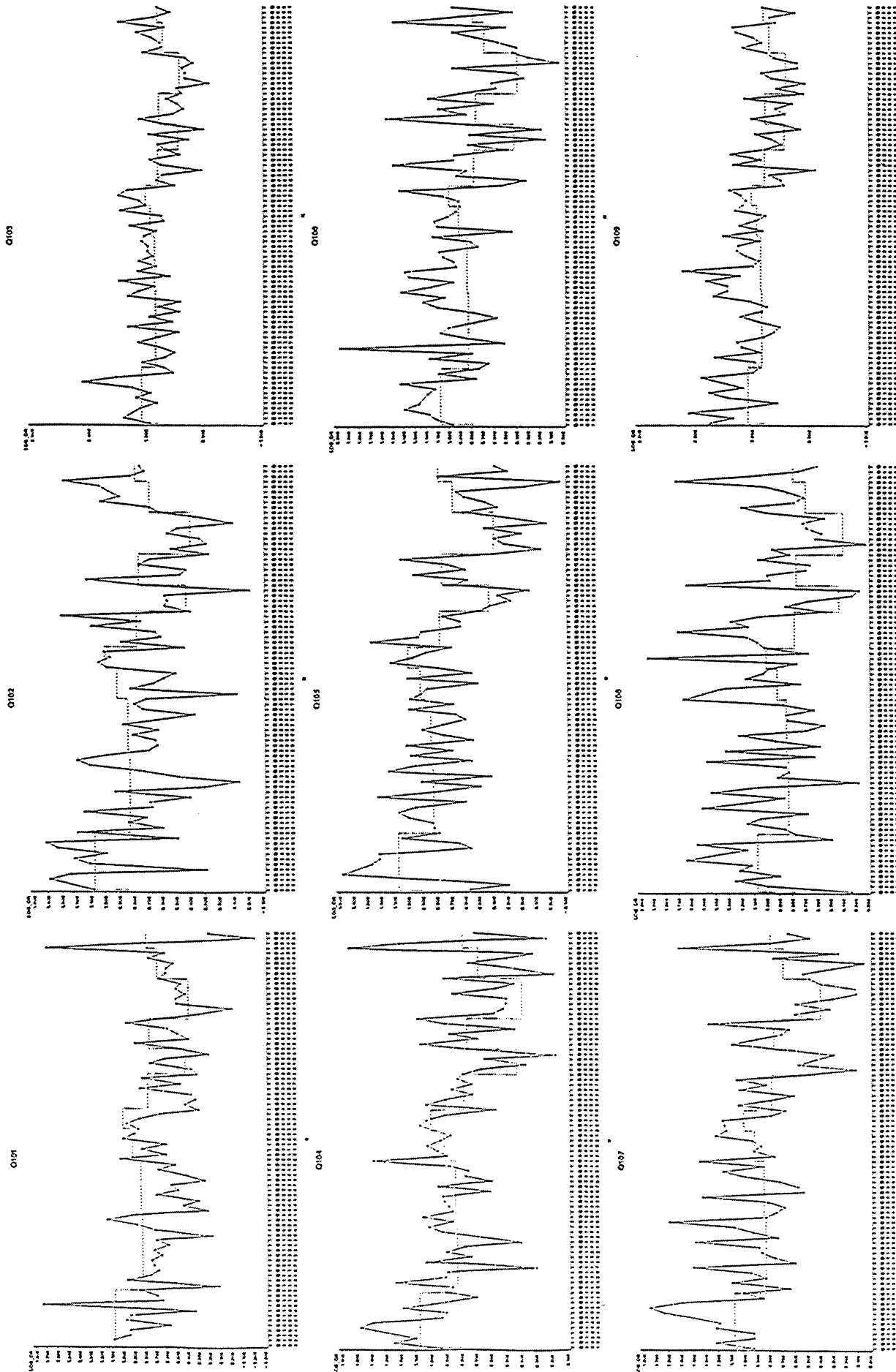


Tableau 9. Graphique des log(OR)  
entre les items Q401 à Q408 et les autres items

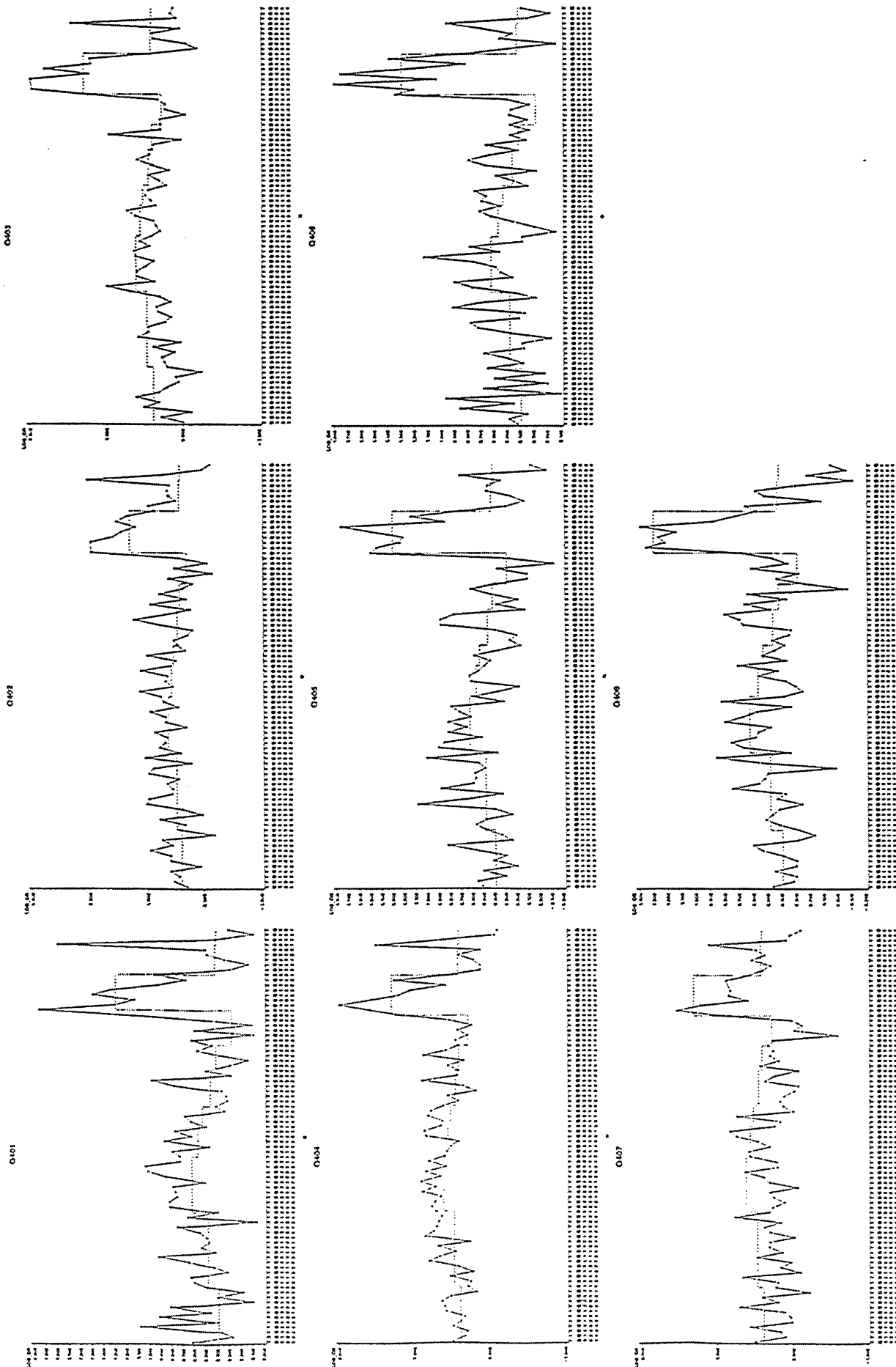
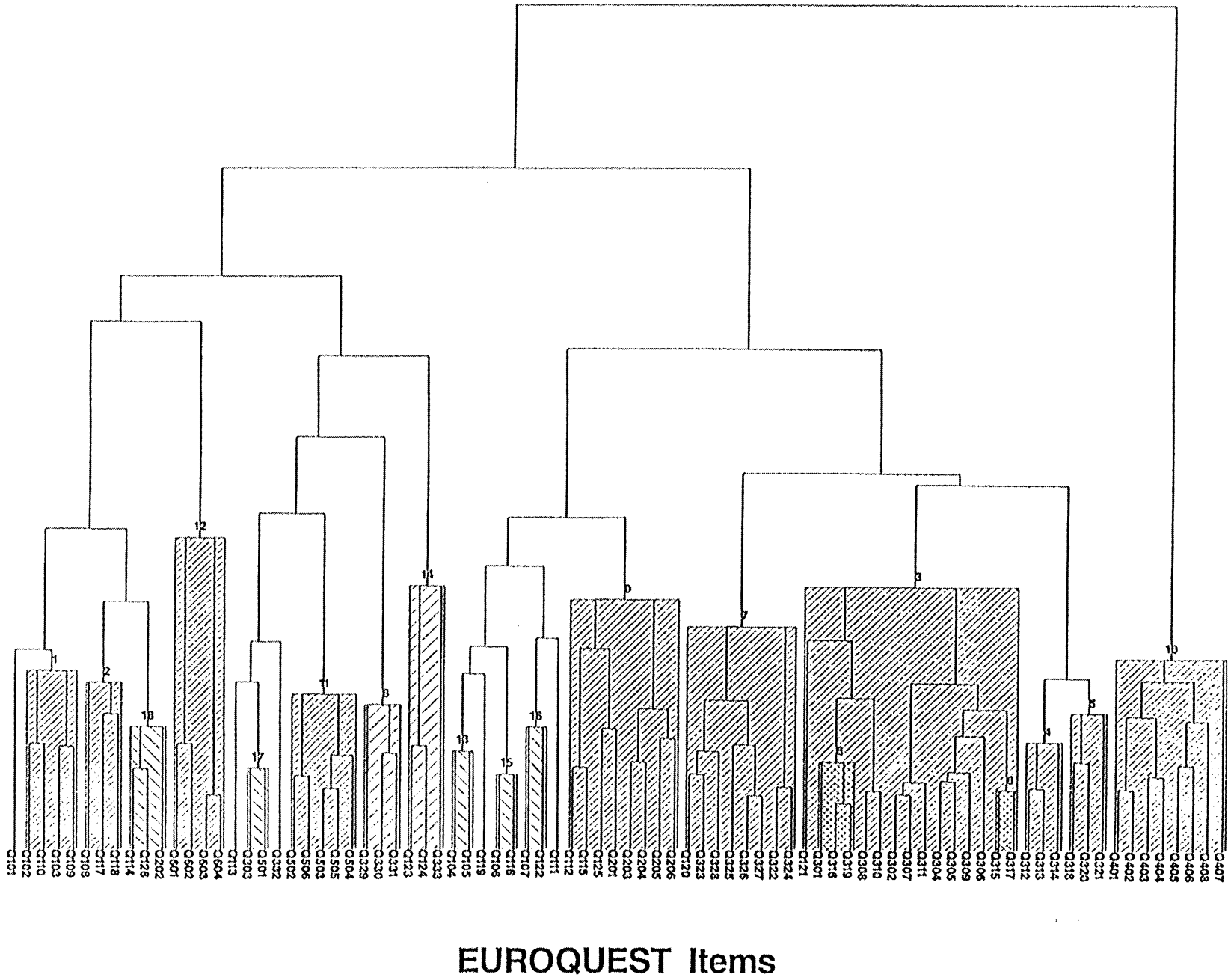


Tableau 10. Classification hiérarchique ascendante



## Tableau 11. Nouveaux domaines d'EUROQUEST

- 1 - fatigue
- 2 - symptômes aigus I
- 3.1 - modification de l'humeur\*
- 3.2 - troubles de concentration\*
- 4 - hypersensibilité
- 5 - anxiété
- 6 - perception de la santé et de la qualité de vie
- 7.1 - troubles de mémoire subjectifs\*
- 7.2 - troubles de mémoire « objectivables » \*
- 8.1 - troubles neurologiques\*
- 8.2 - troubles neurovégétatifs\*
- 9.1 - troubles du sommeil\*
- 9.2 - troubles du transit\*
- 10 - atteintes SNC précoces
- 11 - troubles de l'équilibre
- 12 - troubles du goût
- 13 - personnalité inquiète
- 14 - symptômes aigus II

\* domaines individualisés car ne s'agréant que tardivement dans la classification hiérarchique (le 3.1 au 3.2, le 7.1 au 7.2, etc.).

Tableau 12. Moyenne des logarithmes des Odds Ratios entre deux items appartenant au même domaine ou à l'un des 17 autres domaines

D1	D2																			
	Quest- ion Isolee	Fatig- ue	Aigus I	Abili- te	Intrat- ion	sensi- bilite	Anxie- te V	Anxie- te VI	subje- ctive	objec- tive	logiq- ues	osoma- tiques	Somme- il	Anxie- te	Preco- ces	Equil- ibre	Gout etc	Inqui- ete	Aigus II	ALL
	LOG_OR	LOG_OR	LOG_OR	LOG_OR	LOG_OR	LOG_OR	LOG_OR	LOG_OR	LOG_OR	LOG_OR	LOG_OR	LOG_OR	LOG_OR	LOG_OR	LOG_OR	LOG_OR	LOG_OR	LOG_OR	LOG_OR	LOG_OR
	MEAN	MEAN	MEAN	MEAN	MEAN	MEAN	MEAN	MEAN	MEAN	MEAN	MEAN	MEAN	MEAN	MEAN	MEAN	MEAN	MEAN	MEAN	MEAN	MEAN
Question Isolee	0.545	0.683	0.648	0.671	0.762	0.386	0.581	0.428	0.778	0.715	0.742	0.674	0.453	0.448	0.761	0.804	0.701	0.776	0.727	0.626
Fatigue	0.683	1.349	0.816	0.965	1.090	0.475	0.711	0.762	0.884	0.873	0.762	0.959	0.529	0.595	0.822	0.953	0.790	0.932	1.002	0.844
Symptomes Aigus I	0.648	0.816	1.073	0.750	0.825	0.286	0.446	0.552	0.718	0.571	0.669	0.692	0.403	0.632	0.826	0.879	0.796	0.539	0.999	0.687
Irritabilite	0.671	0.965	0.750	1.545	1.281	0.632	0.770	0.776	0.852	0.919	0.891	0.867	0.544	0.796	0.768	1.025	0.891	1.197	1.014	0.912
Troubles Concentration	0.762	1.090	0.825	1.281	1.619	0.512	0.812	0.903	1.119	1.096	1.070	0.877	0.650	0.531	0.976	1.128	0.893	1.108	0.986	0.950
Hypersensibilite	0.386	0.475	0.286	0.632	0.512	1.290	0.367	0.407	0.480	0.645	0.356	0.588	0.361	0.489	0.447	0.443	0.393	0.668	0.541	0.534
Anxiete V	0.581	0.711	0.446	0.770	0.812	0.367	1.686	0.612	0.730	0.652	0.920	0.790	0.540	0.415	0.441	0.732	0.393	1.192	0.723	0.690
Anxiete VI	0.428	0.762	0.552	0.776	0.903	0.407	0.612	1.577	0.600	0.762	0.826	0.789	0.364	0.537	0.631	0.872	0.574	0.948	0.839	0.691
Memoire subjective	0.778	0.884	0.718	0.852	1.119	0.480	0.730	0.600	1.678	1.393	0.957	1.212	0.643	0.717	0.999	0.785	0.977	0.889	0.936	0.851
Memoire objective	0.715	0.873	0.571	0.919	1.096	0.645	0.652	0.762	1.393	1.497	0.855	0.863	0.428	0.604	0.814	0.899	1.022	0.770	0.886	0.817
Troubles Neurologiques	0.742	0.762	0.669	0.891	1.070	0.356	0.910	0.826	0.957	0.855	1.609	1.020	0.466	0.436	0.841	1.018	1.030	0.944	1.127	0.818
Troubles Psychosomatiques	0.674	0.959	0.692	0.867	0.877	0.588	0.790	0.789	1.212	0.863	1.020	1.349	0.608	0.692	0.810	0.792	0.958	0.848	1.129	0.834
Troubles Sommeil	0.453	0.529	0.403	0.544	0.650	0.361	0.540	0.364	0.643	0.428	0.466	0.608	1.244	0.412	0.499	0.497	0.523	0.747	0.505	0.513
Troubles Sommeil Anxiete	0.448	0.595	0.632	0.796	0.531	0.489	0.415	0.537	0.717	0.604	0.436	0.692	0.412	1.039	0.579	0.612	0.609	0.663	0.770	0.595
Atteintes SNC Precoces	0.761	0.822	0.826	0.768	0.976	0.447	0.441	0.631	0.999	0.814	0.841	0.830	0.499	0.579	1.474	1.154	0.861	0.852	0.851	0.758
Troubles Equilibre	0.804	0.953	0.879	1.025	1.128	0.443	0.732	0.872	0.785	0.899	1.018	0.792	0.497	0.612	1.154	1.972	0.889	0.819	1.216	0.868
Troubles Gout	0.701	0.790	0.796	0.891	0.893	0.393	0.393	0.574	0.977	1.022	1.030	0.958	0.523	0.609	0.861	0.889	1.382	0.719	1.173	0.768
Personnalite Inquiete	0.776	0.932	0.539	1.197	1.108	0.668	1.192	0.948	0.899	0.770	0.944	0.848	0.747	0.663	0.852	0.819	0.719	1.628	0.992	0.887
Symptomes Aigus II	0.727	1.002	0.999	1.014	0.986	0.541	0.723	0.839	0.536	0.886	1.127	1.129	0.505	0.770	0.851	1.216	1.173	0.992	1.730	0.913
Difference between the global score and the domain score		0.51	0.39	0.63	0.67	0.76	1.00	0.89	0.83	0.68	0.79	0.52	0.73	0.44	0.72	1.10	0.61	0.74	0.82	

**ANNEXE : Questionnaire EUROQUEST**  
**Version française**

# QUESTIONNAIRE DE SYMPTOMES NEUROTOXIQUES

EUROQUEST

Version française

Dominique CHOUANIERE\*

Maryvonne FOURNIER\*

Claire GERAUD-MAHUT\*\*

Sarah HILLCOAT-NALLETAMBY\*\*\*

Angela VERDIER\*\*\*\*

\* INRS, Service Epidémiologie, Avenue de Bourgogne, BP 27, 54501 Vandoeuvre Cedex

\*\* Hôpital Fernand Widal, 200 rue du Faubourg St Denis, 75010 Paris

\*\*\* 8 Chemin Scribe, 92190 Meudon

\*\*\*\* Chevalet/Mazeyrat, Aurouze, 43230 Paulhaguet

## EUROQUEST

### CONSIGNES DE REMPLISSAGE

Utilisez pour le remplissage une règle en suivant ligne par ligne le texte

NUMERO D'IDENTIFICATION :

Date de naissance :

TOUS LES RENSEIGNEMENTS QUE VOUS FOURNIREZ SONT CONFIDENTIELS

Nous avons tous, de temps en temps, certains troubles. Ce questionnaire concerne des troubles qui vous seront peut être familiers. Il vous faudra indiquer si vous avez eu ces troubles au cours de ces derniers mois en cochant la case qui convient.

Nous vous invitons à répondre à toutes les questions.

Pour chaque question, vous avez 4 réponses possibles.

Par exemple pour les premières, les réponses possibles sont :

Jamais / très rarement      Quelquefois      Souvent      Très souvent

Si vous avez eu ce trouble assez souvent, cochez la case "Souvent" et ainsi de suite.

Si vous avez des difficultés à choisir entre les réponses, fiez-vous à votre première impression.

I AU COURS DE CES DERNIERS MOIS, AVEZ-VOUS :

(cochez une seule case par question)

	Jamais ou très rarement	Quelquefois	Souvent	Très souvent
1/ laissé tomber involontairement ce que vous aviez dans les mains ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2/ manqué de force dans les bras et dans les jambes ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3/ remarqué une perte de la sensibilité des mains et des pieds ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4/ eu une sensation de lourdeur ou d'engourdissement des membres ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5/ eu des picotements (ou des fourmillements) dans les bras ou les jambes ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6/ eu des problèmes d'équilibre ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7/ eu l'impression que les choses n'ont pas le même goût ou la même odeur ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Au cours de ces derniers mois, avez-vous :

	Jamais ou très rarement	Quelquefois	Souvent	Très souvent
8/ eu l'impression d'avoir la peau du visage endormie ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9/ eu des difficultés à contrôler les mouvements de vos mains ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10/ remarqué une lenteur dans ce que vous avez l'habitude de faire ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11/ eu les mains qui tremblaient ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12/ eu mal à la tête ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13/ transpiré sans raison ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14/ eu des nausées (mal au coeur) ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15/ eu mal au ventre ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16/ eu des vertiges ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17/ été essoufflé(e) sans avoir fait d'efforts physiques ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18/ eu des palpitations au coeur ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
19/ eu des bourdonnements (ou des sifflements) d'oreille ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
20/ eu une impression de grande fatigue ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
21/ ressenti une absence de désir sexuel ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
22/ eu du mal à supporter l'alcool, si vous en prenez ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Au cours de ces derniers mois, avez-vous :

	Jamais ou très rarement	Quelquefois	Souvent	Très souvent
23/ eu des diarrhées ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
24/ été constipé(e) ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
25/ manqué d'appétit ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
26/ eu l'impression d'avoir la tête prise dans un étau (comprimée) ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

II AU COURS DE CES DERNIERS MOIS, PENDANT ET APRES LE TRAVAIL, AVEZ-VOUS EU :  
(cochez une seule case par question)

	Jamais ou très rarement	Quelquefois	Souvent	Très souvent
1/ les yeux irrités (rouges, qui piquent et qui pleurent) ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2/ l'impression d'être ivre (ou saoul(e)) sans avoir bu d'alcool ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3/ la bouche ou la gorge sèche ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4/ la gorge irritée (qui vous donnait envie de tousser tout le temps) ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5/ le nez qui coulait ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6/ un mauvais goût dans la bouche ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

III AU COURS DE CES DERNIERS MOIS, AVEZ-VOUS :

(cochez une seule case par question)

	Jamais ou très rarement	Quelquefois	Souvent	Très souvent
1/ eu des difficultés à vous mettre au travail ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2/ eu l'impression d'être sur les nerfs ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3/ eu des moments de "déprime" ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4/ ressenti de l'impatience ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5/ été contrarié(e) par des choses sans importance ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6/ eu du mal à rester en place ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7/ eu des sautes d'humeur (des changements d'humeur) ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8/ manqué d'intérêt pour ce qui vous entoure ? (sentiment d'indifférence ou de détachement)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9/ manqué d'entrain ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10/ manqué d'intérêt pour les activités qui vous mettent en contact avec les autres ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11/ eu des difficultés à contenir votre colère ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12/ eu tendance à oublier des choses ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13/ eu besoin de noter les choses pour vous en souvenir ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14/ tendance à oublier ce que vous voulez dire ou faire dans la vie de tous les jours ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Au cours de ces derniers mois, avez-vous :

	Jamais ou très rarement	Quelquefois	Souvent	Très souvent
15/ eu du mal à vous concentrer ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16/ eu tendance à être perdu(e) dans vos pensées, (à rêver) ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17/ eu l'esprit embrouillé au moment de vous concentrer ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18/ eu du mal à vous souvenir des noms et des dates ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
19/ tendance à être distrait(e), à être dans la lune ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
20/ eu des difficultés à vous rappeler ce que vous avez lu, ou vu à la télévision ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
21/ entendu dire que vous perdiez la mémoire ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
22/ eu tendance à vous endormir sans que ce soit le moment ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
23/ été anormalement ou excessivement fatigué(e) le soir ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
24/ eu envie de dormir au cours de la journée ? (somnolent(e))	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
25/ été fatigué(e) au réveil ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
26/ manqué d'énergie ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
27/ ressenti une fatigue ou une lassitude générale ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
28/ eu besoin de plus de sommeil qu'auparavant ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**V POUVEZ-VOUS, A PRESENT, INDIQUER SI VOUS ETES D'ACCORD OU PAS AVEC LES PHRASES SUIVANTES EN UTILISANT LES 4 POSSIBILITES DE REPONSES ?**

*(cochez une seule case par question)*

	Pas du tout d'accord	Pas d'accord	D'accord	Tout à fait d'accord
1/ je suis généralement quelqu'un d'inquiet	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2/ j'arrive moins facilement que les autres à surmonter mes difficultés	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3/ je me fais beaucoup de souci pour des choses sans importance	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4/ j'ai souvent l'impression qu'un malheur peut arriver n'importe quand	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5/ je suis souvent débordé(e) par des choses sans importance.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6/ je manque généralement de confiance en moi	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**VI POUR FINIR, POUVEZ-VOUS REpondre A CES 4 QUESTIONS ?**

*(cochez une seule case par question)*

1/ Votre état de santé est-il ?

Très bon	Bon	Mauvais	Très mauvais
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

2/ Et si vous le comparez avec votre état de santé d'il y a 5 ans, est-il aujourd'hui ?

Meilleur	A peu près pareil	Pire	Bien pire
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

3/ En ce moment, comment trouvez-vous la vie en général ?

Bonne	Assez bonne	Pas très bonne	Mauvaise
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

4/ Et si vous la comparez avec votre vie d'il y a 5 ans, est-elle aujourd'hui ?

Bien Meilleure	A peu près pareille	Pire	Bien pire
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

MERCI D'AVOIR BIEN VOULU REpondre  
A CE QUESTIONNAIRE