

Comment nous en sortons-nous en tant qu'internistes?

Erreurs systématiques de raisonnement dans la pratique clinique quotidienne

Levy Jäger^a, Stefan Markun^a

^a Institut für Hausarztmedizin, Universitätsspital Zürich, Zürich

Cet article est un résumé de l'atelier ayant le même intitulé, qui a eu lieu lors du congrès d'automne 2020 de la SSMIG. L'article englobe des bases théoriques, ainsi que les vignettes de cas présentées et leurs solutions, y compris les réponses de 48 médecins ayant participé à l'atelier. Afin que les lecteurs puissent approfondir leur réflexion et s'autotester, les auteurs leur recommandent de répondre aux questions et aux vignettes (lien dans l'encadré sur cette page) avant même de poursuivre la lecture de cet article.

Introduction / contexte

L'erreur est humaine. Il existe néanmoins de nombreuses situations dans lesquelles les erreurs peuvent avoir de lourdes conséquences, en particulier lorsqu'il est question de patients.

Dans la pratique médicale quotidienne, les erreurs ne sont le plus souvent pas dues à un manque de compétence ou encore à une intelligence insuffisante; une explication simple, plausible et pas véritablement décourageante parvient fréquemment à être identifiée. C'est ce qu'ont révélé les résultats des recherches en psychologie cognitive menées au cours de la dernière décennie, moyennant l'utilisation d'un modèle, qui est d'une manière générale capable de très bien décrire nos processus de pensée. D'après cette *théorie du double processus*, les processus décisionnels peuvent être décrits comme une interaction de deux systèmes cognitifs [1]. Les premières conclusions rapides sont tirées par le «système 1», qui est intuitif. Ce système n'est que partiellement soumis à notre contrôle conscient, mais il ne nécessite qu'un faible effort de notre part. Les réflexions plus fines sont, quant à elles, attribuables au «système 2», qui est plutôt lent et analytique; ce dernier nous demande certes plus d'efforts, mais il est accessible de façon plus consciente. C'est grâce au système 1 que, dès le premier contact avec un patient, nous identifions déjà certains schémas qui nous conduisent d'emblée à un diagnostic de travail. Cela est certainement d'une grande utilité dans les situations d'urgence.

Les questions et vignettes de cas présentées lors de l'atelier peuvent être consultées par les lecteurs via le lien suivant: <https://de.surveymonkey.com/r/ihamz2020>.



Toutefois, ces conclusions hâtives peuvent également nous induire en erreur. C'est là qu'entre en jeu le système 2, qui peut moduler et éventuellement corriger les décisions prises par le système 1 via des réflexions rationnelles. Cela ne signifie aucunement que le système 1 veut uniquement nous orienter sur de fausses pistes. Bien au contraire: c'est grâce à lui que nous parvenons à gérer la plupart des situations du quotidien sans efforts trop importants. Il met à notre disposition des heuristiques simples, que nous pouvons appliquer avec succès à une grande partie de nos décisions.

Dans l'idéal, nous devrions toujours prendre nos décisions via un calibrage des deux systèmes. Toutefois, nos réserves d'énergie ne sont pas toujours suffisantes pour impliquer le système 2, qui demande assez d'efforts. Le quotidien clinique accaparant, les conflits avec les collègues ou un manque de sommeil peuvent déjà suffire à mettre notre appareil logique hors service. Les conclusions sont alors moins soumises à une modulation par le système 2, et il y a un risque accru d'erreurs. En d'autres termes: les erreurs se produisent souvent non pas par incompetence, mais bien plus parce que les réserves d'énergie de notre appareil de pensée sont limitées.

En nous appuyant sur les questions et vignettes de cas présentées lors de l'atelier, nous présenterons différentes erreurs de raisonnement qui se produisent fréquemment dans le quotidien clinique. Nous souhaitons cependant tout d'abord définir deux catégories de base:

- Un sophisme (*fallacy*) désigne toute utilisation d'une conclusion invalide dans une chaîne de raisonnement. Nous nous concentrons ici sur les sophismes qui n'ont pas de lien direct avec le contenu des arguments utilisés dans ce contexte. Nous souhaitons tout particulièrement aborder les sophismes qui sont associés à des estimations et interprétations erronées de la fréquence de certains événements.
- Un **biais cognitif** (*bias*) désigne une erreur systématique dans les processus de pensée et de décision, qui échappe aux règles de la pensée rationnelle.

Depuis l'introduction de ces termes par les travaux révolutionnaires des psychologues Tversky et Kahneman, de nombreux exemples les illustrant ont pu être identifiés dans les domaines les plus variés [2]. Les erreurs de raisonnement dans la médecine clinique font désormais l'objet de recherches actives et sont régulièrement présentées dans de vastes articles de revue (par ex. [3]).

Questions et vignettes

1. Exercices d'échauffement (CRT)

Les questions 1–3 sont des éléments du *Cognitive Reflection Test* (CRT) [4]. Le CRT évalue la propension des volontaires à mener une réflexion consciente (système 2) à propos de leur réponse intuitive rapide (système 1).

Question 1.1

«Une raquette et une balle coûtent à elles deux 110 francs suisses. La raquette coûte 100 francs de plus que la balle. Combien coûte la balle?»

Face à cette question, le système 1 fait rapidement la différence entre 110 et 100 et il fait s'imposer l'idée que la balle doit coûter 10 francs. Le système 2 recalcule et constate la chose suivante: si la balle coûte 10 francs, la raquette coûte alors 90 francs de plus que la balle et non pas 100 francs de plus, comme mentionné dans l'énoncé. La solution correcte, dans laquelle à la fois la somme de 110 francs et la différence de 100 francs sont respectées, est: balle = 5 francs, raquette = 105 francs.

La réponse correcte a été donnée par 48% des participants.

Question 1.2

«Si 5 machines ont besoin de 5 minutes pour fabriquer 5 objets, combien de minutes faut-il à 100 machines pour fabriquer 100 objets?»

Face à cette question, le système 1 reconnaît les proportions constantes entre le nombre de machines, le nombre d'objets et la durée de fabrication (5,5,5) et propose une solution analogue (100,100,100). Le système 2 reconnaît néanmoins que la durée de fabrication n'augmente pas si le nombre de machines augmente dans la même proportion que les objets à produire. La solution correcte est donc: 100 machines ont également besoin de 5 minutes pour fabriquer 100 objets (100,100,5).

La réponse correcte a été donnée par 73% des participants.

Question 1.3

«Une nappe de nénuphars se trouve à la surface d'un lac. Chaque jour, la nappe double de taille. Si 48 jours sont nécessaires pour que la nappe recouvre l'ensemble du lac, combien de temps faut-il pour que la moitié du lac soit recouverte?»

Le système 1 réagit au mot «double» ainsi qu'à la consigne de calcul rétroactif avec une division par deux et il propose $48/2=24$ comme solution. Le système 2 constate que le doublement est quotidien et que la moitié du lac a donc dû être recouverte au Jour $48-1=47$.

La réponse correcte a été donnée par 69% des participants.

2. Auto-évaluation et comparaison avec le CRT

«A quel point êtes-vous doué pour éviter de commettre des erreurs de raisonnement (auto-évaluation par rapport à la moyenne des participants de l'atelier)?» Les participants avaient le choix entre cinq quintiles allant de «parmi les 20% les meilleurs» jusqu'à «parmi les 20% les plus mauvais».

Dans l'hypothèse d'une auto-évaluation parfaite par rapport aux autres, les participants devraient être répartis en groupes de taille égale sur la base de leurs auto-évaluations. En réalité, 46% ont estimé qu'ils se situaient «parmi les 20% moyens». Par la suite, nous avons comparé les auto-évaluations des différents participants avec le nombre de réponses correctes (points) obtenues au CRT (fig. 1). Sur la base du résultat obtenu au CRT, il s'est avéré que 31% des participants se sont sous-estimés et que 10% se sont surestimés.

3. Vignettes cliniques

Question 3.1

«Dans un grand cabinet médical urbain, un programme de dépistage de la BPCO est en place depuis dix ans pour les fumeurs et anciens fumeurs âgés de >60 ans ayant (eu) une consommation de >40 paquets-années. Sur la base d'une longue expérience, on sait qu'une BPCO peut être mise en évidence chez 50% de tous les patients dépistés. Lors du dépistage des quatre der-

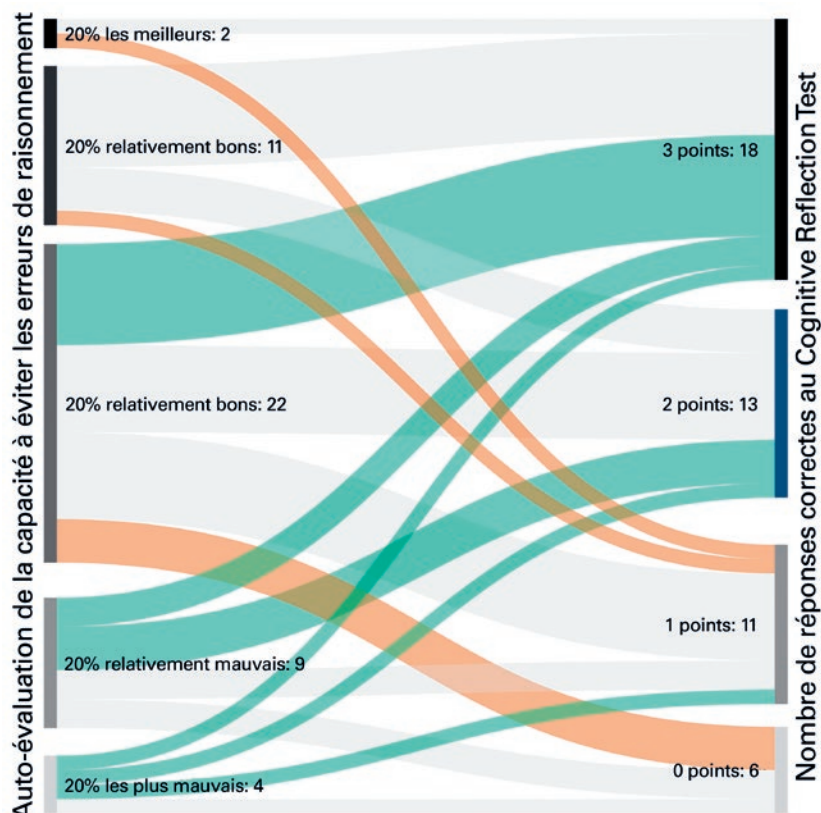


Figure 1: Diagramme de Sankey des réponses à la question «A quel point êtes-vous doué pour éviter de commettre des erreurs de raisonnement?» (proportions représentées par la hauteur des barres verticales gauches), et comparaison avec le nombre de réponses correctes au *Cognitive Reflection Test* (proportions représentées par la hauteur des barres verticales droites). La largeur des jonctions entre les barres gauches et les barres droites est proportionnelle au nombre de participants avec des combinaisons de réponses correspondantes (les jonctions rouges correspondent à une surestimation de soi, les jonctions vertes correspondent à une sous-estimation de soi).

niers patients, une BPCO a été objectivée quatre fois de suite. A combien estimez-vous la probabilité de BPCO chez le prochain patient dépisté?»

Le système 1 simplifie la situation et part intuitivement du principe d'un nombre total raisonnable de patients, à partir duquel un échantillonnage est réalisé. En cas de faible nombre total de patients, le pool de patients atteints de BPCO serait en effet épuisable et il serait possible de déduire les probabilités futures de diagnostic sur la base des diagnostics préalablement posés. Dans le cas présenté ci-dessus, il est tentant de supposer que la probabilité de BPCO chez le patient suivant est inférieure à 50%. Dans la réalité clinique, le pool de patients avec un diagnostic donné est cependant très grand et il est impossible d'émettre des pronostics fiables pour les futurs patients sur la base de quelques observations ou d'une accumulation de cas sur le court terme. Dans le cas ci-dessus, l'expérience de longue

date montre que la probabilité est donc toujours de 50%. Le sophisme qui est fait ici s'appelle «*gambler's fallacy*» (sophisme du joueur) et il s'apparente au joueur de roulette qui, après qu'une couleur soit sortie plusieurs fois, mise sur l'autre couleur pour supposément augmenter sa chance de gagner.

Lors de l'atelier, 85% des participants ont répondu correctement à cette vignette.

Question 3.2

«Une femme de 72 ans est amenée dans un cabinet médical par sa fille, car le matin même, elle ne savait brièvement plus en quelle année nous étions. En supposant une déshydratation, la fille a fait boire une grande quantité de liquide à sa mère, suite à quoi cette dernière était à nouveau orientée dans le temps. Toutefois, une incontinence urinaire a été constatée. L'assistante médicale réalise aussitôt un test urinaire par bandelette, qui montre une leucocyturie et est positif pour les nitrites. La température corporelle, la pression artérielle et le pouls de la patiente sont normaux. La patiente est actuellement parfaitement orientée (temps, espace, personne, situation). Les loges rénales ne présentent pas de douleurs à la percussion. La patiente a une infection urinaire. Après l'exclusion d'une inflammation pertinente (CRP normale et hémogramme sans particularités), un traitement ambulatoire par antibiotiques avec un contrôle de suivi d'ici un à trois jours est suffisant. Etes-vous d'accord avec cette approche?»

Cette vignette contient des pièges à deux niveaux. Au premier niveau, la patiente se présente déjà chez l'assistante médicale avec un tableau suggestif. Le système 1 dispose déjà de résultats qui plaident en faveur d'une infection urinaire. Le système 1 tente alors encore de générer d'autres résultats qui soutiennent l'hypothèse (test urinaire par bandelette positif), sans toutefois collecter d'autres résultats qui remettraient en question l'hypothèse (par ex. test pour rechercher une hypercalcémie). Cette prise en compte sélective d'informations confirmant l'hypothèse aux dépens d'informations la réfutant s'appelle «*confirmation bias*» (biais de confirmation). Au deuxième niveau, le médecin voit la patiente avec un diagnostic pré-tracé. Les examens supplémentaires servent uniquement encore à déterminer la sévérité du diagnostic pré-tracé; des diagnostics alternatifs ne sont pas envisagés. Le fait de se cantonner à une piste déterminée sur la base d'informations (erronées) préalables s'appelle «*anchoring*» (ancrage).

Dans l'atelier, 42% des participants ont approuvé l'approche décrite dans cette vignette (antibiotiques en partant du principe d'une infection urinaire, sans recherche d'une autre cause malgré la confusion). La construction de cette vignette ne permet toutefois pas de distinguer de façon certaine une omission involontaire (sophisme) et une omission volontaire (par ex. en raison d'une probabilité pré-test trop faible) d'examens supplémentaires.

Question 3.3

«Une patiente de 25 ans a pris rendez-vous à votre cabinet en raison d'une fatigue fluctuante depuis plusieurs années. Elle est une défenseuse engagée des droits des animaux, est végétarienne et fait beaucoup de sport. Qu'est-ce qui est le plus probable chez cette patiente: 'Elle a une anémie' ou 'Elle a une anémie avec une carence en fer'?»

Cette vignette contient des facteurs de risque de carence en fer. Le système 1 s'accapare immédiatement de ce stéréotype et néglige le fait que les anémies ferriprives ne constituent qu'un sous-groupe parmi toutes les anémies possibles. La globalité des anémies non autrement spécifiées est donc plus probable que le sous-groupe des anémies ferriprives. Le fait de miser à tort sur un stéréotype spécifique et non pas sur un cas moins spécifique est appelé «*conjunction fallacy*» (erreur de conjonction).

Lors de l'atelier, 38% des participants ont répondu correctement à cette vignette.

Question 3.4

«Chez la patiente de la Question 3.3, des examens suffisants ne révèlent ni anémie (Hb 13,5 g/dl), ni carence en fer (ferritine 64 ng/ml), ni autres causes objectivables de fatigue. La patiente avait déjà passé des examens il y a un an qui avaient livré des résultats similaires, avec notamment une ferritine de 70 ng/ml, et aucune mesure spécifique n'avait été initiée. Toutefois, comme la fatigue était particulièrement prononcée, la patiente a consulté de sa propre initiative un spécialiste en thérapie ferrique. Ce dernier lui a administré une perfusion de fer, suite à quoi la fatigue s'est nettement améliorée. La patiente est convaincue que cela suggère clairement que la perfusion de fer a agi sur la fatigue. Partagez-vous cet avis?»

Il est impossible de déduire un lien de causalité certain à partir de l'auto-observation non contrôlée de la patiente. D'autres causes pourraient expliquer l'amélioration des symptômes, comme par exemple l'évolution naturelle ou un effet placebo. La supposition d'un lien de causalité en raison de la séquence temporelle est un sophisme appelé «*post hoc ergo propter hoc*».

Lors de l'atelier, 71% des participants n'ont pas partagé l'avis de la patiente.

Conclusion

Les observations faites lors de l'atelier nous montrent que les participants semblaient avoir une forme mentale suffisante pour mobiliser leur système 2. Bien entendu, il est cependant impossible d'en tirer des conclusions quant aux réserves cognitives générales des internistes: un biais de sélection entre par exemple en ligne de compte, vu que ce sont les médecins particulièrement intéressés et préalablement informés qui étaient inscrits à notre atelier. Fait intéressant, il y avait net-

tement plus de participants qui se sont sous-estimés par rapport aux autres participants que de participants qui se sont surestimés. La vignette pour laquelle les participants ont eu les plus mauvais résultats était celle traitant de l'erreur de conjonction. Cela pourrait suggérer que les participants pourraient être enclins à se laisser induire en erreur par des stéréotypes, conduisant à poser à la hâte un diagnostic précis alors que cela n'est pas justifié. Les médecins s'en sont le mieux sortis pour la vignette traitant du sophisme *post hoc ergo propter hoc*. Cela suggère que la plupart des participants estiment à juste titre qu'il faut s'appuyer sur des observations contrôlées pour évaluer l'efficacité d'un traitement.

Enfin, il reste à nous demander: Pouvons-nous faire quelque chose pour éviter de telles erreurs de raisonnement? Ou alors devons-nous nous faire à l'idée que certaines erreurs restent inévitables pour nous? La situation n'est fort heureusement pas désespérée et différentes techniques, qui sont regroupées sous le terme *debiasing* et sont censées aider à éviter les biais et les sophismes, ont été décrites. Les preuves quant à leur efficacité sont certes limitées, mais il vaut parfaitement la peine de se pencher dessus [5]. La technique du *cognitive forcing* («forçage cognitif») s'est avérée efficace pour la pratique clinique quotidienne [3]. Elle consiste à freiner volontairement son propre processus de pensée et à méditer dessus, autrement dit à pratiquer la métacognition. Par exemple, il peut déjà être utile de faire une petite pause avant la pose d'un diagnostic et de s'assurer que l'on a suffisamment considéré et activement réfuté tous les diagnostics différentiels alternatifs envisageables du symptôme principal, et en particulier les plus dangereux d'entre eux.

Ces approches reposent sur l'espoir que le simple fait de connaître et d'avoir conscience de telles erreurs peut aider à éviter d'en commettre de nouvelles. Observer de façon répétitive les sophismes et biais commis par nous-mêmes et par les autres dans le quotidien (pas seulement) clinique permet sans doute d'acquérir une plus grande sensibilité vis-à-vis de cette thématique. Et qui sait, peut-être que le seul fait d'avoir participé à notre atelier ou d'avoir lu cet article permettra d'éviter de commettre l'une ou l'autre erreur.

Références

- 1 Crokerry P. Clinical cognition and diagnostic error: applications of a dual process model of reasoning. *Adv Health Sci Educ Theory Pract*. 2009;14(Suppl 1):27–35.
- 2 Tversky A, Kahneman D. Judgment under Uncertainty: Heuristics and Biases. 1974;185(4157):1124–31.
- 3 O'Sullivan ED, Schofield SJ. Cognitive bias in clinical medicine. *J R Coll Physicians Edinb*. 2018;48(3):225–32.
- 4 Frederick S. Cognitive Reflection and Decision Making. *Journal of Economic Perspectives*. 2005;19(4):25–42.
- 5 Ludolph R, Schulz PJ. Debiasing Health-Related Judgments and Decision Making: A Systematic Review. *Med Decis Making*. 2018;38(1):3–13.

Correspondance:
Dr. med. Stefan Markun
Institut für Hausarzt-
medizin
Universitätsspital Zürich
Pestalozzistr. 24
CH-8091 Zürich
stefan.markun[at]usz.ch