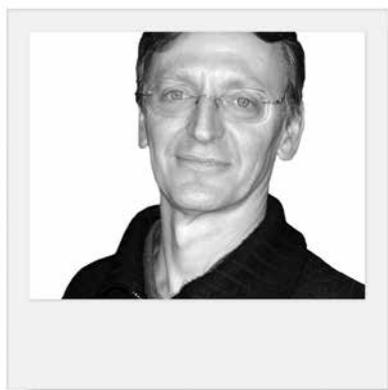


Festina lente

*Hâtez-vous lentement et, sans perdre courage,
Vingt fois sur le métier remettez votre ouvrage.*

~ Boileau



→ F. DIÉVERT
Clinique Villette, DUNKERQUE.

Lors de la lecture de la publication d'un essai thérapeutique contrôlé, il est parfois difficile de comprendre où peut se situer un éventuel biais qui atténue, voire parfois invalide, les conclusions mises en avant par les auteurs. Cela requiert du temps, une méthode et une certaine expérience reposant sur un esprit critique. Alors, pensez donc, quelles conclusions fiables tirer au sortir d'une présentation dans un grand congrès suite à la visualisation en 12 minutes d'une quinzaine de diapositives présentant les résultats d'un essai thérapeutique contrôlé? Présentation elle-même faite au milieu de trois à quatre autres relatant des études souvent très différentes sur des sujets eux aussi différents. Et c'est ainsi que l'on peut ressortir d'une telle présentation avec des idées fausses, ou plutôt faussées, voire influencées (à dessein?) par les présentateurs.

En ce sens, certains, le plus souvent sans même s'en rendre compte, exploitent les biais cognitifs du spectateur. Un biais cognitif est une erreur de raisonnement qui prédispose notre pensée à adopter un point de vue plutôt qu'un autre. C'est souvent le résultat d'utilisation de raccourcis de la pensée, appelés heuristiques. De tels raccourcis nous permettent de prendre des décisions rapides concernant des problèmes néanmoins complexes en suivant des règles approximatives et instinctives.

La méthode scientifique a été développée, parmi d'autres raisons, pour contrer ces biais et obtenir des connaissances objectives. Le problème est que la présentation d'un travail reposant sur une méthode scientifique ne peut que faire appel à des biais cognitifs. En effet, une présentation est un procédé de communication et, plus elle sera courte, plus elle devra résumer un propos parfois complexe et aura recours à des "raccourcis de pensée": elle reposera donc bien sur l'exploitation des biais cognitifs du présentateur et du spectateur.

Or, en matière de science, et plus encore si celle-ci a pour objet la santé des patients, de nombreux problèmes devraient être examinés avec attention et recul. Mais, au terme d'une présentation courte, nous cédon par facilité à nos heuristiques et finissons souvent par agir en fonction de nos biais. Aussi, lors de commentaires "à chaud" des résultats d'une étude, les implications qui peuvent en être proposées sont parfois non adaptées.

L'historien Fernand Braudel a formalisé le concept de durée dans l'analyse des faits historiques comme suit: "*en surface, une histoire événementielle, qui s'inscrit dans le temps court [...]; à mi-pente, une histoire conjoncturelle, qui suit un rythme plus lent [...]; en profondeur, une histoire, de longue durée, qui met en cause des siècles*". Une présentation scientifique dans un congrès fait indéniablement appel à la notion de temps court, celle de l'événementiel, du commentaire

BILLET DU MOIS

journalistique à chaud. Or, il semble préférable en matière de science et de prise de décision en matière de santé d'avoir recours à un temps un peu plus long, sans pour autant entrer dans la longue durée, inappropriée au problème envisagé. En d'autres termes, dans l'analyse d'une présentation à caractère scientifique, il est souvent utile de se hâter lentement : "*festina lente*".

Notre propos, dans ce billet, est d'illustrer cette approche par trois exemples, requérant trois durées d'analyse différentes. Dans le premier, issu de l'étude REDUCE LAP présentée à l'AHA 2016, l'immédiateté peut primer tant le biais est évident. Dans le deuxième, issu de l'étude DANISH présentée à l'ESC 2016, la réflexion devrait être aussi rapide, mais le type de présentation, en quelques minutes, a pu faire passer une donnée non valide pour un acquis de la science. Enfin, le dernier exemple, issu de l'étude SPRINT présentée à l'AHA 2015, est typiquement, malgré l'ampleur des effets d'annonce accompagnant cette étude, un cas où

une réflexion longue et une analyse détaillée des données sont nécessaires.

L'étude REDUCE LAP : créer une communication interauriculaire est-il bénéfique dans l'insuffisance cardiaque ?

Petite étude, certes, 64 patients inclus (mais le piège n'est pas là), l'étude REDUCE LAP évaluait à 1 an le devenir de patients ayant une insuffisance cardiaque à fraction d'éjection préservée (ICFEP) et ayant eu un dispositif implantable permettant de créer et de laisser ouverte une communication interauriculaire.

En sortant de la présentation des résultats de l'étude, on pouvait croire – et certains ont été jusqu'à l'affirmer – que ce dispositif améliorait le pronostic de l'ICFEP. Toutes les données présentées allaient dans le bon sens et le mode de présentation pouvait induire un tel raccourci de pensée. La **figure 1** représente la 14^e diapositive parmi

les 16 présentées dont 7 ont été consacrées au titre, aux conflits d'intérêts et au contexte de l'étude (épidémiologie, nature du dispositif implantable utilisé...), 8 à la méthode et aux résultats et 1 à la conclusion.

Nulle surprise, donc, à recevoir une revue en ligne le jour même de la présentation annonçant "*L'étude REDUCE LAP-HF : un dispositif de shunt interatrial implanté par voie percutanée montre un bénéfice à 1 an chez les patients ayant une ICFEP*". Nulle surprise non plus à ce que de nombreux commentaires écrits ou filmés en quasi-direct reprennent ces mêmes conclusions. Tout incitait à penser cela dans la présentation : des diagrammes en barre avec les données à l'inclusion, à 6 mois et à 1 an dans lesquels, avec le temps, certaines barres avaient "significativement" évolué dans le bon sens (qualité de vie, distance parcourue au test de marche de 6 minutes, temps d'exercice, débit cardiaque...) et la conclusion des auteurs en quelques phrases :

- le bénéfice clinique et hémodynamique observé à 6 mois s'est maintenu à 1 an, sans preuve d'éventuels effets adverses;
- il y a eu une amélioration clinique pertinente de la classe NYHA, de la capacité d'exercice et de la qualité de vie;
- il y a eu une réduction cliniquement pertinente de la pression capillaire pulmonaire;
- des essais thérapeutiques contrôlés sont maintenant nécessaires et en cours pour déterminer la valeur de cette nouvelle stratégie de prise en charge de l'ICFEP.

Ici, "le piège" est simple : comme il n'y a pas de groupe contrôlé, il est impossible d'évaluer, dans ce qui est observé, ce qui revient au dispositif et ce qui revient à l'évolution naturelle de la maladie et/ou à un effet placebo et/ou à un effet de régression vers la moyenne.

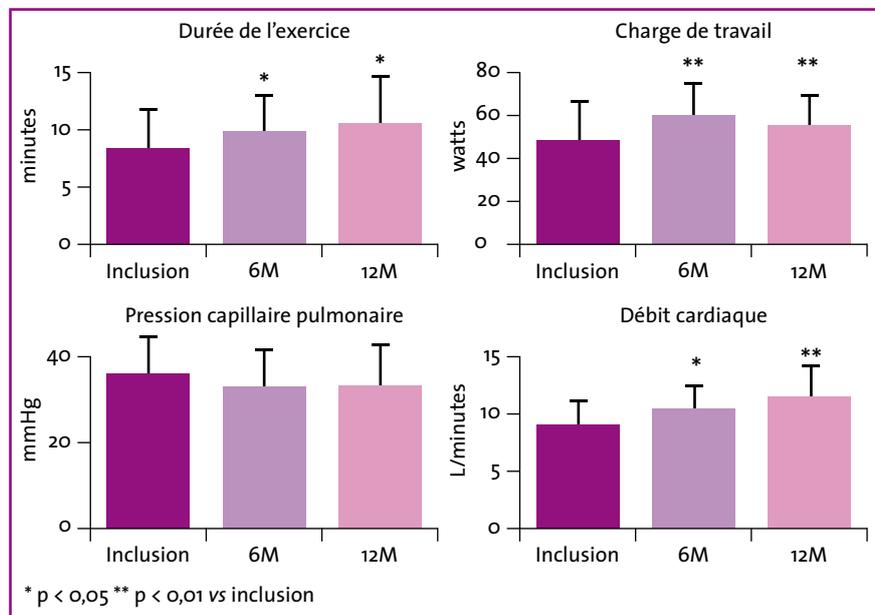


FIG. 1 : Une des diapositives de la présentation des résultats à 1 an de l'étude REDUCE LAP lors du congrès 2016 de l'AHA.

Les renseignements tirés et attendus de cette étude ne sont donc pas des inférences concernant l'effet qu'exerce le dispositif, mais des données brutes (absolues) d'effets secondaires lors de l'implantation et lors du suivi (comme d'éventuelles infections, embolies, thrombus) et les critères indiquant que le dispositif est toujours perméable à 1 an.

Certes, il est absolument nécessaire d'avoir une évaluation prospective de l'ensemble des autres critères présentés, mais celle-ci a pu laisser croire que ce qui était observé était un effet spécifique du dispositif. Or, rien ne permet de l'indiquer, en dehors de la réalisation d'un essai contrôlé avec un groupe témoin ayant une procédure simulée et l'autre une pose effective de l'appareil et en observant s'il y a des différences d'évolution entre les deux groupes.

Les conclusions à tirer de cette étude et de cette présentation semblent être de deux ordres.

>>> La première concerne le fond : les données présentées montrent que l'implantation puis le suivi à 1 an d'un dispositif permettant d'obtenir une *shunt* interauriculaire sont associés à un taux de X % de complications et que le taux de perméabilité du dispositif est de Y % à 1 an. Mais, à la sortie de la présentation, qui a pu dire quels étaient ces deux taux ?

>>> La seconde est sur la forme : quelle terminologie adopter pour qualifier ce qui est observé en termes de paramètres cliniques et/ou hémodynamiques dans l'année de suivi ? Autrement dit, quelles expressions faut-il éviter afin de ne pas faire croire que ce qui est observé est induit par le dispositif ? De toute évidence, le mot "bénéfique" ne devrait pas être utilisé et ce pour deux raisons. La première parce qu'il s'agit d'un terme normatif et non scientifique ; il s'agit d'un jugement de valeur. La seconde est qu'en constatant que, par rapport à l'état de base plusieurs paramètres

évoluent vers des valeurs indiquant un meilleur état de santé, rien n'indique que cette évolution est due au dispositif. Il convient aussi d'éviter les expressions du type "le dispositif réduit" "diminue" "est associé à moins"... Les données présentées ne peuvent qu'induire en erreur et devraient avoir comme utilité principale de renseigner sur l'évolution de patients sélectionnés, chez lesquels a été implanté un dispositif particulier, et ce afin de calculer le nombre de sujets nécessaires à inclure dans un futur essai contrôlé évaluant contre procédure simulée ce dispositif.

L'étude DANISH : le défibrillateur est-il réellement bénéfique chez les sujets de moins de 68 ans ?

L'étude DANISH avait comme objectif d'évaluer si l'implantation d'un défibrillateur automatique (DAI) pouvait diminuer la mortalité totale de patients ayant une insuffisance cardiaque non ischémique (FEVG < 35 % ; NYHA II à III). Elle a été conduite en ouvert, de façon randomisée avec un groupe contrôlé n'ayant pas d'implantation de DAI.

Elle n'a pas démontré la validité de l'hypothèse qu'elle évaluait : au terme du suivi moyen de 68 mois, il n'y a pas eu de différence significative dans l'incidence de la mortalité totale entre les groupes. En revanche, il y a eu une différence significative entre les groupes concernant l'incidence de la mort subite en faveur du groupe ayant eu le DAI par rapport au groupe contrôle.

Pour les auteurs, l'analyse en sous-groupe a montré une hétérogénéité de résultats, un bénéfice semblant apparaître lorsque l'âge était plus bas dans une analyse en tertile d'âge (< 59 ans, de 59 à 68 ans et au moins égal à 68 ans). Et, lors de la présentation, l'auteur de l'étude a montré une

diapositive comparant le sous-groupe des patients de moins de 68 ans à celui des patients âgés d'au moins 68 ans (fig. 2) : cette diapositive "montre" qu'il y a une réduction significative de la mortalité totale (critère primaire) chez les patients de moins de 68 ans et pas d'effet significatif chez les patients d'au moins 68 ans.

De ce fait, au sortir de la salle de présentation, plusieurs médecins ont été persuadés et ont indiqué qu'il devenait nécessaire de modifier la pratique et les recommandations pour ne proposer un DAI que chez les patients ayant une insuffisance cardiaque non ischémique âgés de moins de 68 ans.

La question est : l'étude DANISH démontre-t-elle que le DAI diminue la mortalité totale chez les patients de moins de 68 ans ? **La réponse est non, bien évidemment.** Dans cette étude dont le résultat sur le critère principal est "négatif", pour utiliser la terminologie usuelle, 13 catégories de sous-groupes ont été évaluées : il est donc quasiment inéluctable que, dans une voire deux catégories, il y ait un résultat "significatif" en faveur ou défaveur du traitement évalué, et ce par simple effet de hasard. Mais que font les auteurs de l'étude ? De toute évidence, analysant initialement trois sous-groupes en fonction de leurs âges répartis en tertiles, ils créent à posteriori, une nouvelle analyse en sous-groupe en ne prenant cette fois-ci qu'une seule borne d'âge (68 ans) isolée donc, à posteriori. Toute l'analyse biostatistique est là pour démontrer combien est immense la possibilité de conclure à un résultat significatif qui de fait ne l'est pas, et plus encore, cette démonstration théorique est étayée par de très nombreuses publications ayant complètement invalidé les résultats de conclusions issues de sous-groupes.

Où était le piège ? En premier, il fallait comprendre que l'essai étant "négatif",

BILLET DU MOIS

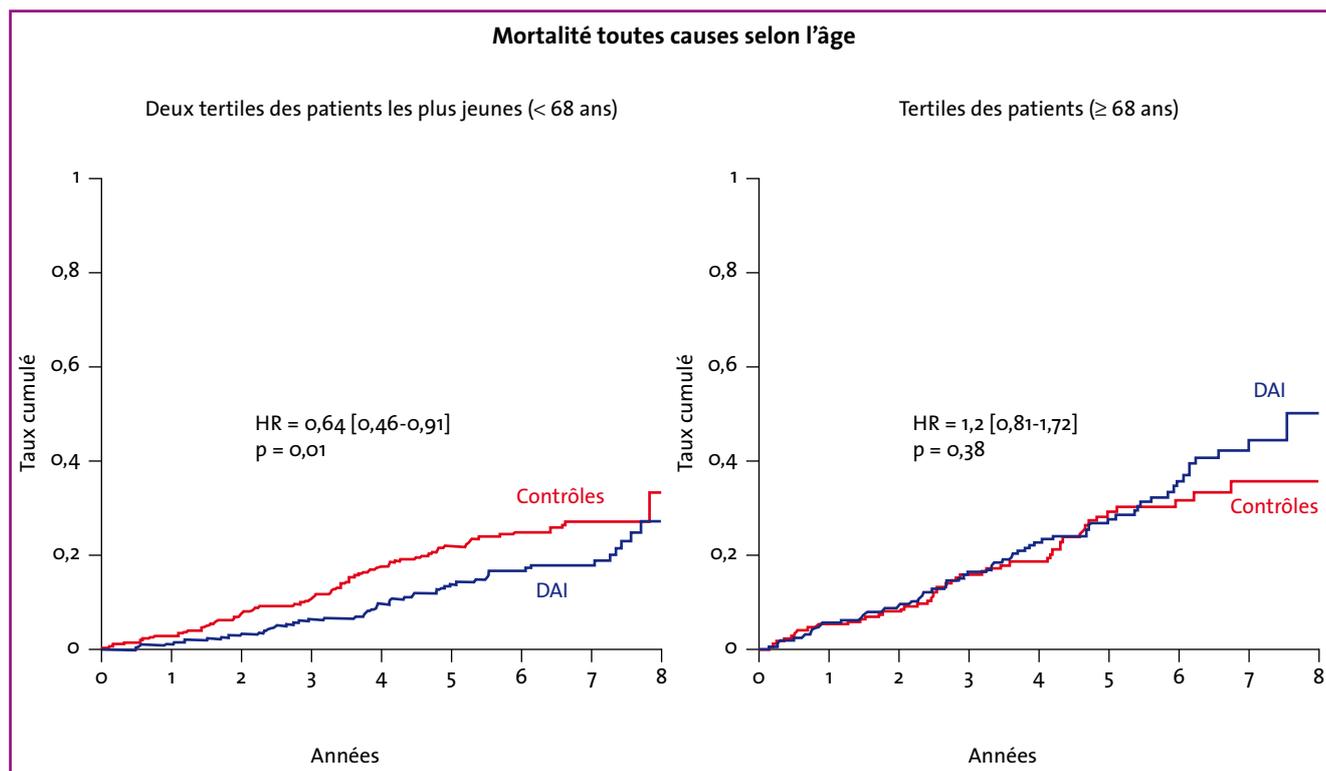


Fig. 2 : Diapositive projetée lors de la présentation des résultats de l'étude DANISH à l'ESC 2016, tendant à montrer un bénéfice chez les patients de moins de 68 ans (sans bénéfice chez les patients d'au moins 68 ans).

les analyses en sous-groupes n'avaient qu'une valeur indicative. Mais surtout, il fallait voir une donnée rapidement présentée, au point de n'avoir souvent pas été vue car l'œil se fixe plus facilement sur les courbes de survie (dans la **figure 2**), que les courbes de survie de la partie de gauche regroupaient celles de deux tertiles de patients et celles de droite, celle du troisième tertile. Oui, c'était écrit sur la diapositive, mais qui se rappelle l'avoir vu d'emblée ? Par ailleurs, nul doute qu'en divisant la population de l'étude en quartiles, quintiles, déciles... centiles... on puisse isoler un sous-groupe dans lequel le DAI augmente la mortalité totale et un sous-groupe dans lequel il la diminue de plus de 50 %...

Dès lors, l'attitude pratique à avoir à l'issue de cet essai est un raisonnement complexe prenant en compte plusieurs éléments parmi lesquels : oui, il y a

une réduction des morts subites avec le DAI mais, non, cette réduction des morts subites ne diminue pas la mortalité totale, et ce probablement parce que l'incidence de la mort subite, comme sa part relative dans la mortalité totale, est très faible. De ce fait, le nombre de patients à traiter pour éviter une mort subite est élevé, un grand nombre de patients n'auront aucun effet de ce dispositif, les exposant inutilement à ses complications et le rapport coût-bénéfice est très élevé, compte tenu du coût du dispositif et de son suivi. Dès lors, il devient nécessaire d'identifier les critères de haut risque de mort subite chez ce type de patients afin de juger de ceux qui peuvent justifier d'une implantation.

Du fait des limites de l'analyse en sous-groupe proposée dans DANISH, il n'est pas sûr que l'âge, avec le seuil de 68 ans notamment, constitue un des mar-

queurs du risque de mort subite et un marqueur de l'utilité du DAI.

L'étude SPRINT : la conclusion globale est-elle valide ? Faut-il diminuer la pression artérielle systolique en dessous de 120 mmHg ?

L'étude SPRINT nous confronte à un sérieux problème. Au départ, il y a un important effet d'annonce, celui d'un résultat "mirifique" annoncé tant dans la presse médicale que grand public au terme d'un communiqué de presse daté de septembre 2015 et indiquant qu'une étude dans l'HTA avait été arrêtée avant son terme pour démonstration d'une réduction de la mortalité totale. Le communiqué émanait du NIH (*National Institute of Health*, aux États-Unis, que l'on pourrait assimiler au ministère de la Santé en France),

car il s'agissait du commanditaire et du financeur de l'étude.

Paradoxe, dans le climat actuel où l'industrie pharmaceutique est accusée de tous les maux, notamment d'influencer les seuils d'intervention (pour les chiffres tensionnels ou lipidiques, par exemple, afin de prescrire plus de traitements), c'est une étude, tout à la fois académique et officielle, qui "démontre" qu'il devient utile de diminuer plus encore les chiffres de pression artérielle, et donc d'avoir recours à plus de traitements antihypertenseurs, privant ainsi les "complotistes patentés" de leur arme de dérision publique préférée (le fameux "tous vendus, tous pourris").

Mais, dès sa présentation, SPRINT, va surprendre nombre d'observateurs car ses résultats sont assez discordants avec ceux de très nombreux essais jusqu'ici disponibles. Surtout, dans les mois qui suivent, les doutes sur la validité de ses conclusions vont être renforcés par la publication de deux études, un essai thérapeutique contrôlé (l'étude HOPE-3) et une analyse d'un registre (CLARIFY) qui vont tous deux renforcer l'hypothèse selon laquelle il existe une courbe en J dans la relation entre les chiffres tensionnels sous traitement et le pronostic: le risque de décès et/ou d'infarctus augmentant de nouveau quand les chiffres tensionnels sont, sous traitement, en dessous de 130 mmHg. **Seule SPRINT indique un résultat inverse.**

Dès lors, depuis la publication des résultats de SPRINT, toutes les données publiées et/ou disponibles la concernant sont progressivement analysées pour essayer de comprendre en quoi son résultat ne peut être considéré comme valide et ne devrait pas modifier la pratique et les recommandations.

1. Retour sur les faits

D'après ses auteurs, dans une conclusion très factuelle de leur étude lors de

sa première présentation à l'AHA 2015, l'étude SPRINT a démontré que le fait de diminuer la pression artérielle systolique (PAS) en dessous de 120 mmHg, plutôt que de la laisser juste en dessous de 140 mmHg, diminue le risque d'événements CV majeurs et notamment la mortalité CV et la mortalité totale, chez des sujets de plus de 50 ans, sans diabète et sans antécédent d'AVC.

Au sortir de la salle de congrès où ont été présentés ces résultats, pour beaucoup de médecins, il n'y avait guère de doutes: il faut diminuer, autant que faire se peut, la PAS en dessous de 120 mmHg. Et pour cause, la mortalité totale diminue avec la stratégie thérapeutique évaluée. Et c'est bien le terme "mortalité totale" qui induit probablement d'emblée un biais cognitif particulier: combien de fois n'a-t-on répété qu'un bénéfice réel ne pouvait s'apprécier que si la mortalité totale était diminuée? Plus encore, cette étude était la grande étude du congrès, annoncée dès avant comme une étude majeure, présentée en session spéciale accompagnée de trois autres présentations expliquant les progrès faits dans l'HTA et l'élément nouveau et fondamental que représentaient les résultats de SPRINT...

Comment, dès lors, garder l'esprit critique et ce d'autant plus que la présentation de la méthode et des résultats est assez simplifiée, car 27 diapositives ont été présentées en 15 minutes, dont celles des remerciements et de la carte des États-Unis où étaient situés les centres investigateurs, laissant peu de possibilité d'en analyser avec précision la méthode?

Enfin, du fait d'un résultat hautement favorable lors d'une analyse intermédiaire, il avait été décidé de terminer l'étude fin août 2015 pour une présentation à l'AHA de novembre 2015 et l'étude a été publiée concomitamment dans le *NEJM* accompagnée de trois éditoriaux (fait rare) et de cinq éditoriaux dans la

revue *Hypertension* (fait rarissime), éléments devant souligner l'importance de ces résultats. **Comment résister à une telle vague, comment développer son esprit critique?**

2. Ce qui a surpris plus ou moins rapidement certains observateurs

Cependant, plusieurs observateurs ont été surpris par la discordance entre les résultats de l'étude SPRINT et ceux des essais antérieurs. Pour certains, cette discordance devait être portée au bénéfice de l'étude SPRINT, signifiant bien qu'il y avait une rupture, une preuve enfin, une avancée majeure. Pour ces commentateurs, les autres essais n'avaient pas réussi à mettre en évidence un résultat qui ne pouvait, intuitivement, être que celui-là: il est bénéfique de diminuer la PAS le plus bas possible. Toute l'épidémiologie d'observation va dans ce sens. On reconnaît là un biais cognitif puissant, celui dénommé "biais de confirmation d'hypothèse": dès lors qu'un résultat est conforme aux attentes préalables, nul n'est besoin d'analyser la validité de la méthode ayant permis de l'obtenir, il s'impose de lui-même comme une réalité.

En revanche, pour ceux qui ont été surpris de l'obtention d'une diminution inhabituelle de mortalité, sans courbe en J dans la relation entre les chiffres tensionnels sous traitement et le pronostic, il devenait urgent d'analyser avec méthode tous les éléments pouvant potentiellement poser le problème de la validité des résultats observés. Et plusieurs pistes ont été explorées à ce jour.

3. La piste de l'insuffisance cardiaque

L'esprit critique remarque que la diminution de mortalité totale observée dans SPRINT, outre le fait d'être exceptionnellement constatée dans un essai dans l'HTA, n'est pas associée à une diminution significative des AVC (ce qui est quasiment la règle dans les

BILLET DU MOIS

essais d'HTA, cet événement étant très "pression-dépendant"), mais est, en première analyse, associée à une diminution significative du risque d'insuffisance cardiaque et à une diminution significative de la mortalité CV. L'inférence a été vite faite par presque tous les observateurs : la stratégie évaluée diminue le risque d'insuffisance cardiaque, ce qui diminue la mortalité CV et donc la mortalité totale. C'est ce que l'on peut du moins supposer à la lecture du **tableau I** (d'après la publication et la présentation de l'étude) présentant les résultats concernant l'ensemble des événements du critère primaire et de ses divers composants (tous IDM, syndrome coronaire aigu non IDM, tous AVC, toutes insuffisances cardiaques, décès CV).

Dès lors, diverses hypothèses ont été avancées pour tenter de comprendre ce résultat. Pour certains, les données épidémiologiques récentes montrant que l'insuffisance cardiaque est devenue l'un des événements principaux survenant chez les sujets hypertendus, il pouvait donc être attendu que cet événement soit fortement et rapidement abaissé par la prise en charge de l'HTA et que cela puisse expliquer le bénéfice. Cette hypothèse est par conséquent en faveur de la validité du résultat.

Pour d'autres, plusieurs éléments constituent plutôt des biais pouvant expliquer la réduction des insuffisances cardiaques et ne permettent donc pas d'affirmer que les résultats constatés sont fiables et valides, c'est-à-dire imputables exclusivement à la réduction de la PAS.

>>> Premier biais potentiel, l'étude a été conduite en ouvert et le diagnostic d'insuffisance cardiaque n'est parfois pas simple, ce qui a pu conduire les médecins investigateurs, favorables à l'attitude thérapeutique évaluée, à surestimer les cas dans le groupe contrôle et à les minorer dans le groupe évalué.

>>> Deuxième biais, à l'inclusion dans l'étude, 33,5 % des patients avaient une PAS en moyenne inférieure à 132 mmHg, et ces patients avaient des traitements antihypertenseurs, notamment des diurétiques, des inhibiteurs de l'enzyme de conversion (IEC) ou des antagonistes des récepteurs de l'angiotensine 2 (ARA2). Une fois randomisée, la moitié de ces patients a probablement eu un renforcement du traitement antihypertenseur afin d'avoir une PAS inférieure à 120 mmHg, mais l'autre moitié... a-t-elle eu une diminution de son traitement, notamment des diurétiques et/ou des IEC et/ou ARA 2 afin de l'amener vers la cible de 140 mmHg,

ce qui a pu potentiellement augmenter son risque d'insuffisance cardiaque ?

>>> Troisième biais potentiel : l'étude étant conduite en ouvert, les patients du groupe devant avoir une PAS inférieure à 120 mmHg pouvaient être suivis avec une plus grande fréquence que ceux du groupe contrôle afin de surveiller leurs chiffres tensionnels et d'adapter le traitement pour obtenir et maintenir une PAS inférieure à 120 mmHg. Dans ce cas, si le patient avait une dyspnée et/ou des œdèmes, n'a-t-il pas eu une majoration des diurétiques permettant de diminuer son risque d'insuffisance cardiaque ? La diminution du risque proviendrait alors non pas de la diminution de la PAS mais d'une différence dans le suivi des deux cohortes.

4. La piste de l'insuffisance cardiaque est-elle la bonne ?

Mais, de fait, à bien y regarder, le **tableau I** ne fournit qu'une explication imparfaite du résultat concernant le critère primaire. Il ne permet pas de savoir si le résultat concernant le critère primaire est bien le fait d'une réduction significative des insuffisances cardiaques sans diminution significative des AVC. Pourquoi ? Parce que le critère primaire prend en

	Groupe intensif		Groupe contrôle		HR (IC 95 %)	Valeur de p
	Nombre d'événements	Taux annuel	Nombre d'événements	Taux annuel		
Critère primaire	243	1,65	319	2,19	0,75 (0,64-0,89)	< 0,001
Tous IDM	97	0,65	116	0,78	0,83 (0,64-1,09)	0,19
SCA non IDM	40	0,27	40	0,27	1,00 (0,64-1,55)	0,99
Tous AVC	62	0,41	70	0,47	0,89 (0,63-1,25)	0,50
Toutes IC	62	0,41	100	0,67	0,62 (0,45-0,84)	0,002
Décès CV	37	0,25	65	0,43	0,57 (0,38-0,85)	0,005

HR : hazard ratio ; IC 95 % : intervalle de confiance à 95 % ; IDM : infarctus du myocarde ; SCA non IDM : syndromes coronaires aigus n'étant pas des infarctus du myocarde ; AVC : accidents vasculaires cérébraux ; IC : insuffisance cardiaque ; CV : cardiovasculaires.

TABEAU I : Résultats de l'étude SPRINT concernant son critère primaire et chacun des éléments de ce critère primaire.

compte la survenue du premier événement parmi ceux en faisant partie, alors que l'analyse critère par critère prend en compte la totalité des événements survenus pendant la durée de l'étude. Cela est une convention de présentation des résultats et est facilement observable dans le tableau : dans le groupe ayant une PAS cible inférieure à 120, il y a eu 243 événements du critère primaire et 298 événements totaux de ce critère, et dans le groupe ayant une PAS cible inférieure à 140, il y a eu 319 événements du critère primaire et 391 événements totaux de ce critère. Ce tableau ne nous renseigne donc pas sur l'incidence réelle des premiers événements cliniques survenus et ne permet pas d'inférence, c'est-à-dire de conclure que c'est bien la réduction des insuffisances cardiaques (en termes de premier événement) et non celle des AVC, par exemple (toujours

en termes de premier événement) qui est à l'origine de la diminution significative des événements du critère primaire. L'observateur à l'esprit critique se trouve donc pour partie démuné.

L'élément le plus marquant de l'étude SPRINT étant la réduction de la mortalité totale, il semble capital de comprendre à quoi peut être due cette réduction, en d'autres termes si certains éléments ne relèvent pas du hasard ou de biais dans la méthode. Si le tableau sur chacun des événements du critère primaire n'est pas informatif, comment faire ? Une nouvelle fois, il faut lire l'appendice publié dans le *NEJM* en même temps que les résultats principaux de l'étude et aller jusqu'à la page 31 afin d'analyser le tableau qui en fournit toutes les causes (**tableau II**). Ici, l'analyse est rendue possible par le fait que, comme la mortalité totale n'était pas le

critère primaire, tous les cas et leurs causes sont référencés.

Il y a donc eu 155 décès dans le groupe intensif et 210 dans le groupe standard. Et, si l'analyse du tableau des composants du critère primaire peut laisser penser que la diminution de la mortalité totale est due à la diminution de la mortalité CV (parce que significative), elle-même due à la diminution de l'insuffisance cardiaque (parce que significative), l'analyse du tableau des causes de mortalité conduit à d'autres conclusions. En effet, il y a eu 8 décès par insuffisance cardiaque dans le groupe intensif et 9 dans le groupe standard. Première conclusion : ce n'est pas la diminution de l'insuffisance cardiaque qui explique la diminution de mortalité totale. Ce n'est pas non plus la diminution de mortalité par AVC (8 cas dans le groupe intensif et 9 cas dans le groupe

	Total	Groupe intensif	Groupe contrôle
Cardiovasculaire	102	37	65
Coronaire	50	18	32
AVC	17	8	9
Mort subite	13	2	11
Insuffisance cardiaque	17	8	9
Non cardiaque mais cardiovasculaire	5	1	4
Non cardiovasculaire	192	90	102
Insuffisance rénale	2	1	1
En rapport avec une dialyse	1	0	1
Cardiaque non ischémique	2	0	2
Cancer	101	49	52
Accident/suicide/homicide	14	4	10
Non cardiaque, non AVC	72	36	36
Non déterminée	71	28	43
Non classifiable	35	13	22
Pas encore attribuée	36	15	21
Total des décès	365	155	210

TABLEAU II : Causes des décès dans l'étude SPRINT.

BILLET DU MOIS

standard). Par ailleurs, trois éléments semblent se détacher : le premier est un moindre nombre de décès coronariens (18 vs 32), le deuxième, un moindre nombre de décès par mort subite (2 vs 11) et le troisième, qui perturbe l'interprétation, un moindre nombre de décès de cause non déterminée dans le groupe intensif (n = 28) que dans le groupe standard (n = 43).

On remarque alors que la différence du nombre de décès CV entre les groupes est de 28 cas et que celle des décès non CV (qui comprend une part de décès de cause indéterminée) est de 27 cas. Deuxième conclusion possible, la mortalité totale est donc autant diminuée du fait de la réduction des décès CV que des décès non CV. Cela peut rendre compte d'un effet hasard ou chance dans le résultat constaté car il est difficile d'envisager que la diminution de la PAS en dessous de 120 mmHg ait un effet favorable sur la mortalité non CV.

Par ailleurs, à l'analyse du diagramme de suivi des patients analysés pour enrôlement, puis randomisés, il est possible de constater que parmi les 9361 patients randomisés, il y a eu 245 perdus de vue (2,6 %), et ce en seulement 3 ans. Cela a deux implications. La première est de faire douter de la qualité du suivi de l'étude. La deuxième est de faire douter de la validité de son résultat car il y a eu plus de perdus de vue que le nombre d'événements permettant d'obtenir une différence significative entre les groupes comparés sur la plupart des critères, notamment sur le critère primaire et sur la mortalité totale. Ainsi, ce taux de perdus de vue ne permet pas de considérer comme parfaitement fiables les résultats enregistrés sur plusieurs des critères majeurs.

5. Et... le détail qui tue l'étude

En mai 2016, SE Kjeldsen (avec trois coauteurs) faisait paraître dans la revue *Hypertension* un article qui tendait à

remettre complètement en cause les résultats de l'étude SPRINT. Il a exposé les tenants et aboutissants du raisonnement qu'il a défendu lors de l'ESC 2016 à Rome et en décembre 2016, à Paris, au cours d'une lecture qu'il a été invité à donner durant l'inauguration des 36^{es} Journées de l'HTA.

En substance, ce médecin norvégien a remarqué que, dans le protocole de l'étude, il était précisé que la PAS des patients, à l'inclusion et pour la surveillance, était enregistrée par un appareil automatique, après 5 minutes de repos en dehors de la présence du médecin, ce qui apparente la mesure à celle obtenue en automesure et non en clinique. De ce fait, une valeur moyenne de PAS en dessous de 120 mmHg, telle qu'elle est mesurée dans l'étude SPRINT, ne correspond pas à une même valeur en mesure clinique, elle lui est inférieure. Pour confirmer que ce mode de mesure fournit des valeurs très basses, il a utilisé deux types de données indirectes.

>>> Le premier type est la valeur des données de mesure ambulatoire de pression artérielle (MAPA) dans une étude complémentaire de SPRINT : or, la valeur de la PAS à 27 mois est à 120 mmHg en mesure dite clinique et à 123 mmHg en moyenne sur 24 heures en MAPA. C'est un paradoxe que la mesure clinique soit inférieure à la valeur obtenue en MAPA, indiquant donc que la mesure clinique a suivi un protocole inhabituel.

>>> Le deuxième type est celui des données d'études ayant évalué la valeur correspondante d'une mesure au repos sans la présence du médecin et en mesure clinique réelle, c'est-à-dire en présence du médecin. L'auteur en a conclu que la valeur de PAS à 120 mmHg fournie pour le groupe prise en charge intensive de l'étude SPRINT correspond probablement à une mesure située entre 130 et 140 mmHg en réelle mesure clinique et que celle fournie comme proche de 140 mmHg dans SPRINT correspond

probablement à une mesure située entre 150 et 160 mmHg en réelle mesure clinique.

De ce fait, il est probable que, si le résultat de l'étude SPRINT est fiable, il n'indique pas qu'il faille diminuer la PAS de 140 mmHg à 120 mmHg en pratique clinique, mais simplement de plus de 150 mmHg à moins de 140 mmHg, ce qui devient compatible avec les données jusqu'ici acquises, à l'exception du bénéfice enregistré sur la mortalité totale, qui ne laisse de surprendre.

SE Kjeldsen va par ailleurs plus loin dans son argumentation. D'une part, il pense que l'élément qu'il a mis en évidence constitue un biais très gênant pour la validation des implications de l'étude SPRINT. D'autre part, ce qui le conforte dans ce caractère gênant est qu'il n'arrive pas à obtenir les détails exacts de la mesure de la PAS par les investigateurs de l'étude SPRINT et que le NIH a aussi des conflits d'intérêts. Pour lui, le NIH est une administration qui tire ses subsides de l'État et a intérêt à afficher des résultats devant permettre de garantir l'allocation de futurs budgets.

6. En synthèse

L'étude SPRINT a un résultat qui n'est pas concordant avec ceux de multiples essais effectués dans l'HTA. Cette discordance en elle-même n'invalide pas le résultat mais conduit à rechercher ce qui peut l'expliquer.

Deux grandes voies sont explorées.

>>> La première a pour principe d'envisager que le résultat n'est pas scientifiquement valide. Il serait consécutif soit à un voire des biais dans la conduite de l'étude : étude ouverte, suivi différent des patients, diminution des traitements chez des patients dont la PAS était basse à l'inclusion... Il pourrait aussi être le fait du hasard favorisé par divers biais :

