

Addiction aux jeux de hasard et d'argent en ligne : comment mettre en place une prévention basée sur les preuves grâce aux Big Data ?



Amandine LUQUIENS

Psychiatre addictologue, Hôpital Paul Brousse, Villejuif

L'addiction aux jeux d'argent en ligne est devenue un grave problème de santé publique. Pour y faire face, il est possible d'utiliser les données laissées par les joueurs sur les sites de jeu. Ces données permettent de repérer l'apparition de comportements à risque, et de proposer aux intéressés des outils favorisant un jeu responsable. Peut-être faut-il, pour bénéficier d'une prévention efficace, accepter les inconvénients d'une surveillance systématisée.

Big Data et prévention

L'engouement sociétal actuel pour le Big Data n'épargne pas la médecine, où l'on oscille entre fascination, fantasmes, vraies avancées et perspectives. Il faut tout d'abord discerner la e-santé et la digitalisation. La e-santé correspond à l'intégration de technologies d'information et de communication dans les soins. La digitalisation correspond aux processus sous-jacents d'innovation et de changement, en particulier le développement de logiciels et de matériel visant in fine l'aide à la décision médicale¹. La masse des données exploitables en médecine est en augmentation exponentielle comme dans les autres secteurs. Différents types de données déjà utilisées en médecine ont été recensés, qu'elles viennent du dossier patient et du soin ou d'autres sources : (a) données biologiques, (b) données d'environnement, (c) dossiers médicaux dématérialisés, (d) enregistrements via des objets connectés (exemples : géo localisation, mais aussi enregistrements d'équipements médicaux comme ceux d'un service de réanimation) (e) publications sur les réseaux sociaux, (f) données issues des téléphones portables². La diversité des données disponibles permet donc aussi bien de suivre à distance le rythme cardiaque d'un patient portant un pace maker, que de dépister un épisode dépressif à travers la publication de posts sur Twitter. Pour ce dernier exemple, il s'agit donc de dépister de manière indirecte un trouble mental chez un individu ne l'ayant pas forcément sollicité, à partir d'un versant de son comportement, les échanges sur un réseau social. D'autres données issues d'un comportement sur internet permettent de faire de la prévention. La pratique des jeux de hasard et d'argent en ligne en fait partie. Mais la nature des données disponibles constitue une configuration exceptionnelle : les données permettent de décrire précisément et dans son intégralité un comportement potentiellement problématique. On a donc non pas une reconstitution indirecte du comportement, mais sa lecture directe.

1. Gehring H., Rackebrandt K. & Imhoff M. 2018. E-Health and reality - what are we facing in patient care? Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz.

Keasberry J., Scott, I. A., Sullivan, C., Staib, A. & Ashby, R. 2017. Going digital: a narrative overview of the clinical and organisational impacts of eHealth technologies in hospital practice. Aust Health Rev, 41, 646-664.

2. Mooney, S. J. & Pejaver, V. 2017. Big Data in Public Health: Terminology, Machine Learning, and Privacy. Annu Rev Public Health.

Données de compte joueur et prévention du jeu problématique

La pratique des jeux de hasard et d'argent en ligne en France est régulée par l'Autorité de régulation des jeux en ligne (ARJEL)³, qui demande aux opérateurs de jeu l'enregistrement de nombreuses données de jeu issues des comptes joueurs. L'ARJEL stocke elle-même ces données sur un serveur frontal⁴, et est autorisée à les utiliser dans une optique de protection du joueur et de prévention de l'addiction depuis 2016. Par ailleurs, ces données peuvent être utilisées par chaque opérateur de jeu dans une optique également de protection du joueur et de prévention de l'addiction. Elles constituent l'opportunité de dépister les joueurs problématiques et de réaliser de la prévention ciblée. En effet, si la prévalence du jeu problématique est en hausse et a été estimée en 2014 à 1,9 % de la population générale française de 15 à 75 ans⁵, l'accès aux soins pour ces personnes est très médiocre. Des initiatives de prévention primaire (éviter l'entrée dans la maladie) et secondaire (favoriser l'accès aux soins et limiter les dommages liés à la maladie) sont donc nécessaires. Deux outils utilisant des algorithmes ont été développés dans le passé pour proposer aux joueurs de repérer un jeu problématique notamment à travers un changement dans leur comportement de jeu. Le plus utilisé, Playscann, a été développé à partir de données de comptes joueurs, mais en toute opacité, sans validation par rapport à une échelle classique de dépistage clinique. Un premier outil de dépistage des joueurs problématiques de poker en ligne a été développé et validé par notre équipe à partir d'un modèle statistique de régression expliquant le score clinique dans l'index canadien du jeu excessif (ICJE) recueilli en ligne, par les données de comportement de jeu enregistrées en routine par l'opérateur de jeu Winamax⁶. Nous avons alors travaillé sur des données agrégées sur le mois. Cette initiative a permis de montrer la faisabilité d'un repérage fiable à partir des données de compte joueur seules, chez des joueurs de poker d'un opérateur. Elle a abouti à une étude fille au sein de l'ARJEL portant sur l'ensemble des activités de jeu en ligne sur les sites français. Cette étude réalisée en 2016 n'a pourtant pas encore eu de traduction en mesure pratique de prévention, probablement du fait d'un paradigme complexe entre liberté de jouer et défiance face au spectre de la surveillance d'une part, argument parapluie porté par les enjeux économiques, et nécessité de mettre en place la prévention ciblée face à une maladie lourde en terme de poids socio-sanitaire, d'autre part.

Au-delà du repérage

Si la mise en place du repérage est complexe car soulevant des questionnements de société, les données de compte joueur permettent aussi de mieux appréhender le processus du jeu problématique, et de mesurer l'efficacité de mesures existantes de jeu responsable, c'est à dire visant à limiter les risques et les dommages des jeux de hasard et d'argent. Plusieurs outils dits de jeu responsables peuvent être utilisés pour aider un joueur à ajuster son comportement de jeu en cas de pratique problématique. Un des outils proposés aux joueurs est l'auto-exclusion, temporaire ou « définitive » (trois ans), ne concernant qu'un opérateur de jeu, ou l'interdiction volontaire de jeu (trois ans également, mais concernant tous les opérateurs de jeu en ligne régulés par l'ARJEL, et les casinos terrestres). Les données disponibles dans la littérature sur les joueurs en ligne s'étant auto-exclus ont montré des résultats peu robustes du fait de problèmes méthodologiques : la majorité des auteurs recrutaient des joueurs volontaires pour participer à l'étude et non l'intégralité des joueurs s'étant auto-exclus, et les interrogeaient *a posteriori* sur leurs motivations. Plusieurs auteurs se sont accordés à dire que les joueurs

3. Assemblée Nationale 2010. Loi n° 2010-476 du 12 mai 2010 relative à l'ouverture à la concurrence et à la régulation du secteur des jeux d'argent et de hasard en ligne- Article 26. In: Sénat, A. N. E. (ed.) NOR: BCFX0904879L.

4. Machine qui reçoit les connexions des utilisateurs

5. Costes J-M., Eroukmanoff V., Richard J-B., Tovar M-L. (2015) Les jeux d'argent et de hasard en France en 2014 *Les notes de l'observatoire des jeux* n°6. Sont classés « problématiques » les joueurs « à risque modéré » (1,5 %) et les joueurs « excessifs » (0,4 %).

6. Luquiens A., Tanguy M. L., Benyamina A., Lagadec M., Aubin H. J. & Reynaud M. (2016). Tracking online poker problem gamblers with player account-based gambling data only. *Int J Methods Psychiatr Res*, 25, 333-342.

s'étant auto- exclus avaient majoritairement une pratique de jeu plus intense que les autres joueurs en ligne sur le même site⁷, mais d'autres ont décrit que seulement une partie d'entre eux présentait un jeu problématique, exploré *a posteriori*⁸. Aucune donnée dans la littérature n'a documenté les motifs d'auto- exclusion au moment de sa mise en œuvre, permettant de lever le biais de rappel et les modifications de comportement à posteriori. A la différence d'études antérieures, et dans la lignée de partenariats précédents entre académiques et opérateurs de jeu (Bwin et « the transparency project »), notre équipe, constituée de professionnels de santé chercheurs (APHP-Inserm) et de mathématiciens du Centre de Mathématiques Appliquées de l'Ecole Polytechnique, travaille sur les données de compte joueur grâce à un partenariat formalisé par une convention laissant clairement la liberté de publication sans droit de regard de l'opérateur de jeu. Les données à notre disposition documentent à la fois les motifs d'auto-exclusion, permettant de repérer les joueurs s'auto-excluant pour un motif évoquant un jeu problématique, et les données de jeu de poker avec une forte granularité, c'est-à-dire un niveau de détail fin de l'information présentée. Le travail sur des données à forte granularité soulève pourtant des questions méthodologiques. En effet, nous sommes face à des mesures répétées dans le temps un grand nombre de fois. D'autre part, le matériau même du jeu est complexe car il est en partie dû au hasard, rendant les variables financières très dispersées : même un joueur pathologique peut gagner sur une main « chanceuse ». La question est donc grossièrement : à quel zoom travailler ? Y a-t-il un vrai intérêt à travailler avec une granularité fine sur des données mettant en jeu en partie le hasard ? Comment gérer les variables à très fort écart type ? Comment prendre en compte l'historique du joueur dans son comportement, à savoir que chaque action de jeu est faite par le joueur en connaissance des précédentes, et que ses prises de décisions sollicitent le fonctionnement cérébral et émotionnel ? Le machine learning se distinguerait de l'approche statistique par le fait qu'il générerait des hypothèses, plutôt que de chercher à les tester. Pourtant, si la masse des données à disposition semble ouvrir ici une boîte de Pandore, le garde-fou revient, comme dans les autres domaines de la médecine et des approches statistiques classiques, à connaître la nature des variables, savoir comment elles ont été obtenues, ce qu'on veut en faire, c'est-à-dire quelle est la finalité pratique de l'analyse.

Enjeux éthiques

En effet, la question « que veut-on faire des données ? » se heurte nécessairement aux enjeux éthiques soulevés par l'exploitation de données issues de comportements en ligne. La balance entre le risque de violation de la vie privée et les intérêts publics doit être faite, et la démarche de l'étude doit nécessairement viser à des bénéfices sociétaux directs ou rapidement applicables⁹. Dans notre cas, plusieurs questions sont soulevées. L'aboutissement des recherches doit-il être un dépistage systématique ou la possibilité de proposer systématiquement un dépistage aux joueurs ? Mais dans ce deuxième cas, à qui et quand faut-il proposer un dépistage ? La faisabilité méthodologique et la création d'un algorithme engage-t-elle éthiquement à son utilisation si elle peut permettre de sauver des vies, sachant le sur-risque suicidaire porté par les joueurs pathologiques ? Posée ainsi, la question reçoit une réponse positive évidente à première vue, qui rencontre le spectre de la « surveillance ». Dans le cas qui nous concerne, la surveillance à visée de prévention est déjà une obligation imposée aux opérateurs de jeu, qui doivent, tous les ans, documenter leurs initiatives de jeu responsable auprès de l'ARJEL dans un rapport dédié, nécessaire pour renouveler leur agrément ; à ceci près que la surveillance est aujourd'hui artisanale, et donc quasi inefficace (quelques dizaines de cas par ans repérés pour certains opérateurs). Toute forme de veille à visée de prévention utilise et collige les données

7. Dragicevic S., Percy C., Kudic A. & Parke J. 2015. A descriptive analysis of demographic and behavioral data from Internet gamblers and those who self-exclude from online gambling platforms. *J Gambl Stud*, 31, 105-32.

8. Griffiths M. D., Wood R. T. & Parke J. 2009. Social responsibility tools in online gambling: a survey of attitudes and behavior among Internet gamblers. *Cyberpsychol Behav*, 12, 413-21.

9. De Lusignan S., Liyanage H., Di Iorio C. T., Chan T. & Liaw S. T. 2016. Using routinely collected health data for surveillance, quality improvement and research: Framework and key questions to assess ethics, privacy and data access. *J Innov Health Inform*, 22, 426-32.

Mittelstadt B. D. & Floridi L. 2016. The Ethics of Big Data: Current and Foreseeable Issues in Biomedical Contexts. *Sci Eng Ethics*, 22, 303-41.

issues des prises en charge individuelles, pour avoir un tableau global, épidémiologique. Ainsi, le pic épidémique de grippe est-il surveillé chaque année en relevant le nombre de cas dépistés et pris en charge. La « surveillance » systématisée est ainsi prévention systématisée si elle est au service de la santé. Paradoxalement, dans le cadre légal actuel, le risque n'est pas à l'utilisation par les opérateurs de ce type d'algorithme à mauvais escient, mais plutôt à leur réticence à le mettre en place, pour ne pas détourner les joueurs de leur pratique. Le deuxième risque est plutôt un risque par omission, c'est-à-dire la non mise en place par les autorités de régulation ou de santé, possiblement pour le même motif économique, alors que l'État bénéficie largement des recettes des jeux. Enfin, l'exemple de notre partenariat, comme d'autres à l'étranger, illustre la faisabilité d'un partenariat académique-privé sur l'exploitation de données. La nécessité d'une convention qui protège les académiques sur la publication sans réserve des résultats s'impose. Il apparaît recommandable d'éviter tout financement direct par l'opérateur afin d'éviter une position de conflit d'intérêt. Dans un domaine où l'État lui-même est en conflit d'intérêt, les professionnels de santé devraient pouvoir se positionner en sentinelles innovantes afin de mobiliser l'État sur des initiatives de prévention dont ils auront démontré la faisabilité et l'efficacité.