

# Le phénotypage digital pour une pratique clinique en santé mentale mieux informée

## Improved Mental Health Clinical Practice Informed by Digital Phenotyping

Alan Bougeard, Rose Guay Hottin, Valérie Houde, Thierry Jean, Thibault Piront, Stéphane Potvin, Paquito Bernard, Valérie Tourjman, Luigi De Benedictis and Pierre Orban

Volume 46, Number 1, Spring 2021

Santé mentale et nouvelles technologies

URI: <https://id.erudit.org/iderudit/1081513ar>

DOI: <https://doi.org/10.7202/1081513ar>

[See table of contents](#)

Publisher(s)

Revue Santé mentale au Québec

ISSN

0383-6320 (print)

1708-3923 (digital)

[Explore this journal](#)

Cite this article

Bougeard, A., Guay Hottin, R., Houde, V., Jean, T., Piront, T., Potvin, S., Bernard, P., Tourjman, V., De Benedictis, L. & Orban, P. (2021). Le phénotypage digital pour une pratique clinique en santé mentale mieux informée. *Santé mentale au Québec*, 46(1), 135–156. <https://doi.org/10.7202/1081513ar>

Article abstract

**Objectives** This review is motivated by the observation that clinical decision-making in mental health is limited by the nature of the measures obtained in conventional clinical interviews and the difficulty for clinicians to make accurate predictions about their patients' future mental states. Our objective is to offer a representative overview of the potential of digital phenotyping coupled with machine learning to address this limitation, while highlighting its own current weaknesses.

**Methods** Through a non-systematic narrative review of the literature, we identify the technological developments that make it possible to quantify, moment by moment and in ecologically valid settings, the human phenotype in various psychiatric populations using the smartphone. Relevant work is also selected in order to determine the usefulness and limitations of machine learning to guide predictions and clinical decision-making. Finally, the literature is explored to assess current barriers to the adoption of such tools.

**Results** Although emerging from a recent field of research, a large body of work already highlights the value of measurements extracted from smartphone sensors in characterizing the human phenotype in behavioral, cognitive, emotional and social spheres that are all impacted by mental disorders. Machine learning permits useful and accurate clinical predictions based on such measures, but suffers from a lack of interpretability that will hamper its use in clinical practice in the near future. Moreover, several barriers identified both on the patient and clinician sides currently hamper the adoption of this type of monitoring and clinical decision support tools.

**Conclusion** Digital phenotyping coupled with machine learning shows great promise for improving clinical practice in mental health. However, the youth of these new technological tools requires a necessary maturation process to be guided by the various concerned actors so that these promises can be fully realized.

# **Le phénotypage digital pour une pratique clinique en santé mentale mieux informée**

**Alan Bougeard<sup>a</sup>**

**Rose Guay Hottin<sup>b 1</sup>**

**Valérie Houde<sup>c</sup>**

**Thierry Jean<sup>d</sup>**

**Thibault Piront<sup>e</sup>**

**Stéphane Potvin<sup>f</sup>**

**Paquito Bernard<sup>g</sup>**

**Valérie Tourjman<sup>h</sup>**

**Luigi De Benedictis<sup>i</sup>**

**Pierre Orban<sup>j</sup>**

- 
1. Alan Bougeard et Rose Guay Hottin ont contribué de façon égale à cet article.
  - a. Étudiant, Centre de recherche de l'Institut universitaire en santé mentale de Montréal.
  - b. Étudiante, Centre de recherche de l'Institut universitaire en santé mentale de Montréal.
  - c. M.D., étudiante, Centre de recherche de l'Institut universitaire en santé mentale de Montréal.
  - d. Étudiant, Centre de recherche de l'Institut universitaire en santé mentale de Montréal.
  - e. Professionnel de recherche, Centre de recherche de l'Institut universitaire en santé mentale de Montréal.
  - f. Ph. D., chercheur, Centre de recherche de l'Institut universitaire en santé mentale de Montréal – professeur sous octroi titulaire, Département de psychiatrie et d'addictologie, Université de Montréal.
  - g. Ph. D., chercheur, Centre de recherche de l'Institut universitaire en santé mentale de Montréal – professeur régulier, Département des sciences de l'activité physique, Université du Québec à Montréal.

**RÉSUMÉ Objectifs** Cette revue trouve sa motivation dans l'observation que la prise de décision clinique en santé mentale est limitée par la nature des mesures typiquement obtenues lors de l'entretien clinique et la difficulté des cliniciens à produire des prédictions justes sur les états mentaux futurs des patients. L'objectif est de présenter un survol représentatif du potentiel du phénotypage digital couplé à l'apprentissage automatique pour répondre à cette limitation, tout en soulignant les faiblesses actuelles.

**Méthode** Au travers d'une revue narrative de la littérature non systématique, nous identifions les avancées technologiques qui permettent de quantifier, instant après instant et dans le milieu de vie naturel, le phénotype humain au moyen du téléphone intelligent dans diverses populations psychiatriques. Des travaux pertinents sont également sélectionnés afin de déterminer l'utilité et les limitations de l'apprentissage automatique pour guider les prédictions et la prise de décision clinique. Finalement, la littérature est explorée pour évaluer les barrières actuelles à l'adoption de tels outils.

**Résultats** Bien qu'émergeant d'un champ de recherche récent, de très nombreux travaux soulignent déjà la valeur des mesures extraites des senseurs du téléphone intelligent pour caractériser le phénotype humain dans les sphères comportementale, cognitive, émotionnelle et sociale, toutes étant affectées par les troubles mentaux. L'apprentissage automatique permet d'utiles et justes prédictions cliniques basées sur ces mesures, mais souffre d'un manque d'interprétabilité qui freinera son emploi prochain dans la pratique clinique. Du reste, plusieurs barrières identifiées tant du côté du patient que du clinicien freinent actuellement l'adoption de ce type d'outils de suivi et d'aide à la décision clinique.

**Conclusion** Le phénotypage digital couplé à l'apprentissage automatique apparaît fort prometteur pour améliorer la pratique clinique en santé mentale. La jeunesse de ces nouveaux outils technologiques requiert cependant un nécessaire processus de maturation qui devra être encadré par les différents acteurs concernés pour que ces promesses puissent être pleinement réalisées.

- 
- h. M.D., psychiatre, Institut universitaire en santé mentale de Montréal – professeure agrégée de clinique, Département de psychiatrie et d'addictologie, Université de Montréal.
  - i. M.D., psychiatre, Institut universitaire en santé mentale de Montréal – professeur adjoint de clinique, Département de psychiatrie et d'addictologie, Université de Montréal.
  - j. Ph. D., chercheur, Centre de recherche de l'Institut universitaire en santé mentale de Montréal – professeur sous octroi adjoint, Département de psychiatrie et d'addictologie, Université de Montréal.

**MOTS CLÉS** apprentissage automatique, apprentissage profond, décision clinique, nouvelles technologies, phénotypage digital, prédiction, pronostic, psychiatrie, psychologie, téléphone intelligent

## Improved Mental Health Clinical Practice Informed by Digital Phenotyping

**ABSTRACT Objectives** This review is motivated by the observation that clinical decision-making in mental health is limited by the nature of the measures obtained in conventional clinical interviews and the difficulty for clinicians to make accurate predictions about their patients' future mental states. Our objective is to offer a representative overview of the potential of digital phenotyping coupled with machine learning to address this limitation, while highlighting its own current weaknesses.

**Methods** Through a non-systematic narrative review of the literature, we identify the technological developments that make it possible to quantify, moment by moment and in ecologically valid settings, the human phenotype in various psychiatric populations using the smartphone. Relevant work is also selected in order to determine the usefulness and limitations of machine learning to guide predictions and clinical decision-making. Finally, the literature is explored to assess current barriers to the adoption of such tools.

**Results** Although emerging from a recent field of research, a large body of work already highlights the value of measurements extracted from smartphone sensors in characterizing the human phenotype in behavioral, cognitive, emotional and social spheres that are all impacted by mental disorders. Machine learning permits useful and accurate clinical predictions based on such measures, but suffers from a lack of interpretability that will hamper its use in clinical practice in the near future. Moreover, several barriers identified both on the patient and clinician sides currently hamper the adoption of this type of monitoring and clinical decision support tools.

**Conclusion** Digital phenotyping coupled with machine learning shows great promise for improving clinical practice in mental health. However, the youth of these new technological tools requires a necessary maturation process to be guided by the various concerned actors so that these promises can be fully realized.

**KEYWORDS** clinical decision, deep learning, digital phenotyping, machine learning, new technology, prediction, prognosis, psychiatry, psychology, smartphone.

## Mesures en santé mentale

Les troubles psychiatriques sont notoirement dynamiques plutôt que statiques dans le temps, suivant des trajectoires temporelles propres à chaque individu (Wright et Woods, 2020). Ces troubles sont fréquemment caractérisés par l’alternance entre des périodes de rémission et des épisodes d’exacerbation marquée des symptômes au fil des mois et années. Les statistiques compilées pour 2018 par l’Institut canadien d’information sur la santé<sup>2</sup> indiquent ainsi que plus de 12 % des patients seront réadmis 2 fois ou plus dans un service hospitalier psychiatrique au Québec dans l’année suivant une première hospitalisation pour cause de trouble mental. À une autre échelle temporelle, les troubles mentaux se caractérisent par des fluctuations plus subtiles de la sévérité des symptômes au fil des jours et semaines (Wichers, 2013). Il apparaît donc critique pour le clinicien en santé mentale de pouvoir suivre de manière suffisamment fine l’évolution de l’état mental des patients afin de prédire adéquatement le risque de rechute et de diminution marquée du fonctionnement ou de la qualité de vie des individus.

La pratique clinique conventionnelle en santé mentale, fondée sur des entretiens en personne dans le cabinet médical hospitalier ou privé, présente certaines limitations dans cette perspective. Premièrement, la surcharge des systèmes de santé limite le nombre de rencontres possibles et le temps allouable à l’évaluation du risque de dégradation de l’état mental futur des patients. Une étude menée en Ontario révèle que les visites reliées à la santé mentale étaient au nombre moyen de 3 par an chez le médecin de famille (13 % de la population provinciale) et de 6 par an chez le médecin psychiatre (3 %) en 2014 (Chiu et coll., 2018). Une autre étude menée auprès d’un groupe multidisciplinaire de 90 praticiens en santé mentale indiquait par ailleurs que moins de 15 minutes par rencontre étaient consacrées en moyenne à l’évaluation du risque (Cohen et coll., 2019). Ces moyennes ne sont évidemment pas le reflet du nombre de rencontres ou du temps alloués aux patients pour lesquels le clinicien juge le risque important. Néanmoins, ces chiffres suggèrent que de potentiels signaux d’alarme ne puissent pas être « mesurés » au moment opportun et ne soient dès lors pas intégrés au corpus d’information dont dispose le clinicien pour guider ses prises de décision clinique de façon personnalisée (Wichers, Schreuder, Goekoop et Groen, 2019 ; Wright et Woods, 2020).

---

2. <https://yourhealthsystem.cih.ca/epub/>

Deuxièmement, la nature même des mesures typiquement obtenues du patient dans le cabinet médical pose question (Fortney et coll., 2017). L'évaluation clinique repose essentiellement sur un jugement subjectif et qualitatif de l'état mental du patient basé sur les symptômes et signes révélés par l'écoute, l'observation et l'interaction. L'usage d'échelles cliniques validées permet une évaluation quantitative de la sévérité des symptômes, du niveau de fonctionnement et de la qualité de vie de façon plus objective. Dans ses directives de pratique pour l'évaluation psychiatrique des adultes, l'Association de psychiatrie américaine appuie l'emploi d'échelles cliniques standardisées, tant pour un entretien initial que pour le suivi de patients en cours de traitement (American Psychiatric Association, 2015). Pourtant, moins de 20% des psychiatres y recourent de façon systématique (Fortney et coll., 2017). Un sondage mené auprès de plus de 600 psychiatres révèle ainsi qu'une minorité de cliniciens utilisent couramment ces outils validés en dépit du fait qu'une majorité d'entre eux jugent que leur prise de décision clinique est améliorée par leur usage (American Psychiatric Association, 2016). Il faut souligner que tant l'évaluation clinique conventionnelle que celle basée sur des échelles quantitatives peuvent souffrir de plusieurs biais cognitifs – biais de rappel, de désirabilité sociale, etc. – qui altèrent la validité de l'information obtenue (Rogler, Mroczek, Fellows et Loftus, 2001).

La pratique clinique en santé mentale repose sur la capacité du clinicien à obtenir les informations pertinentes pour orienter ses prises de décision clinique (figure 1a). L'imperfection de la mesure dans la pratique clinique en santé mentale, en raison de la nature même de l'évaluation ou en raison de la trop faible fréquence des évaluations, pose problème. Idéalement, le clinicien pourrait bénéficier, pour guider sa pratique, de mesures objectives, quantitatives, précises, continues et écologiquement valides, étant acquises dans le milieu de vie du patient. Ces mesures liées aux comportements, aux émotions et à la cognition du patient seraient collectées de façon passive et *in fine* synthétisées afin d'être exploitables dans une perspective clinique (Hirschtritt et Insel, 2018).

## Le phénotypage digital

Un portrait numérique des foyers québécois dressé par le Centre facilitant la recherche et l'innovation dans les organisations (CEFRIO) relève que l'adoption du téléphone intelligent (iPhone, Android, BlackBerry) a

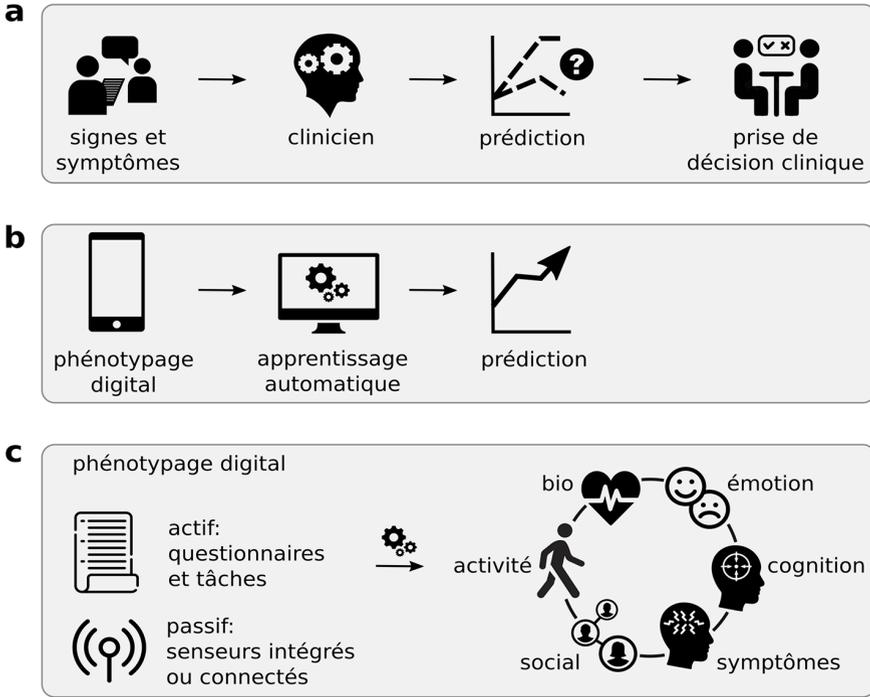
explosé au cours de la dernière décennie au Québec, passant de 13 % en 2009 à 77 % en 2019. Sans surprise, c'est dans les tranches d'âge jeune (18 à 34 ans) que le taux d'adoption le plus élevé (94 %) est observé. Le téléphone intelligent, couteau suisse technologique aux multiples fonctions, est maintenant employé en lieu et place de nombreux appareils appelés à devenir obsolètes: lecteur de musique, appareil photo, navigateur GPS, montre, réveil, calendrier, enregistreur vocal, calculatrice, portefeuille, etc.

L'omniprésence des téléphones intelligents dans les poches de millions d'utilisateurs au Québec offre une fenêtre unique sur les comportements individuels et l'environnement d'une vaste majorité de la population. Au-delà de l'ubiquité des téléphones intelligents, leur usage intensif 24 heures sur 24, 7 jours sur 7 les rend d'autant plus attrayants pour caractériser l'évolution au jour le jour de l'état mental des utilisateurs. Ce processus de quantification, au moyen du téléphone intelligent, du phénotype humain dans son idiosyncrasie, instant après instant et dans le milieu de vie naturel, a été décrit par le terme de phénotypage digital (Onnela, 2021).

Le phénotypage digital repose à la fois sur la contribution active d'informations collectées au travers d'applications mobiles dédiées et sur la collecte passive de flux de mesures générées au travers des multiples senseurs intégrés du téléphone intelligent. Les données dites actives requièrent une participation volontaire, typiquement au travers de la complétion de questionnaires ou de tâches cognitives. Les données dites passives exploitent les nombreux senseurs intégrés au téléphone intelligent (accéléromètre, gyroscope, batterie, processeur, statut de l'écran, capteur de luminosité, GPS, Wi-Fi, Bluetooth, microphone, caméra, clavier, etc.) ainsi que les informations d'exploitation du téléphone (métadonnées des appels téléphoniques, messages texte, applications utilisées, etc.). Une fois passées au travers du moulin de l'apprentissage automatique (Mohr, Zhang et Schueller, 2017), ces différentes données fournissent une panoplie de riches mesures intégratives du phénotype comportemental, émotionnel, social et cognitif (figure 1c). Le phénotypage digital offre ainsi des mesures objectives, quantitatives, continues et écologiquement valides qui répondent à certaines limitations des mesures obtenues dans le cadre de la pratique clinique conventionnelle en santé mentale (Baumeister et Montag, 2019). Au travers d'une revue non systématique, mais représentative de la littérature, nous offrons ici un survol de la richesse

FIGURE 1

Phénotypage digital pour la pratique clinique en santé mentale



a) Illustration du processus conventionnel où la prise de décision clinique repose sur des prédictions imparfaites du clinicien, guidées par les signes et symptômes décelés lors de rares entretiens cliniques avec le patient. b) La prise de décision clinique par le clinicien pourrait être supportée par des prédictions plus justes reposant sur des modèles d'apprentissage automatique incorporant de riches mesures de phénotypage digital, conditionnellement à une interprétabilité suffisante des modèles. c) Le phénotypage digital repose à l'origine sur des données actives volontairement contribuées par l'utilisateur d'une application mobile dédiée ainsi que des données passives densément longitudinales résultant de nombreux capteurs intégrés ou connectés au téléphone intelligent. Des modèles d'apprentissage automatique, potentiellement incorporés à l'application mobile, permettent d'extraire des mesures de phénotypage digital intégratives à partir des mesures densément longitudinales des capteurs, dans différentes sphères pertinentes à la santé mentale.

de ces mesures, avant d'aborder brièvement le rôle de l'apprentissage automatique dans l'aide à la prédiction clinique basée sur ces données de phénotypage digital (figure 1b), et de relever certains enjeux liés à l'adoption des outils proposés tant par les patients que par les cliniciens.

## **Données actives**

### **Autoévaluation**

L'évaluation quantitative de la sévérité des symptômes au travers d'échelles cliniques validées lors de l'entrevue clinique a démontré sa valeur, mais ne remplit pas l'objectif d'une évaluation continue à distance (Goldberg, Buck, Raphaely et Fortney, 2018). Des questionnaires d'autoévaluation couvrant une large panoplie de symptômes psychiatriques ainsi que du niveau de fonctionnement occupationnel ou social et de la qualité de vie ont été adaptés pour faciliter leur administration au travers du téléphone intelligent. Prenant la forme d'échelles de Likert ou d'échelles visuelles analogues, ces questionnaires permettent à l'utilisateur d'autoévaluer ses symptômes de façon journalière, voire plus fréquemment encore au moment même de l'expérience afin de minimiser les failles de reconstruction des expériences à posteriori. En dépit de limitations évidentes liées à la subjectivité des autoévaluations voire plus fondamentalement de la capacité des patients à évaluer tous leurs types de symptômes, de nombreux travaux de recherche ont mis de l'avant la pertinence de ces approches dans les populations souffrant de troubles mentaux (Myin-Germeys et coll., 2018; Wright et Woods, 2020).

### **Tâches cognitives**

Des déficits cognitifs variés (attention, mémoire, fonctions exécutives, etc.) sont observés dans de nombreux troubles mentaux (East-Richard, Mercier, Nadeau et Cellar, 2020). La cognition n'est pas un simple attribut stable d'une personne; elle varie au fil des jours et contextes, covariant avec la motivation, les affects, le sommeil, l'exercice et d'autres facteurs environnementaux et sociaux (Weizenbaum, Torous et Fulford, 2020). Le développement de tests neuropsychologiques adaptés pour le téléphone intelligent fournit des résultats rassurants quant à leurs fidélité et validité (Sliwinski et coll., 2018). Notamment, les données obtenues pour 21 tests cognitifs chez 100 000 participants indiquent d'excellentes propriétés psychométriques pour les adaptations de tests neuropsychologiques standards, avec une fidélité acceptable même pour une courte durée de tâche comprise entre 30 s et 3 min, selon le test (Passell, Dillon, Baker, Vogel et Scheuer, 2019). Des études menées dans des populations psychiatriques parviennent à des conclusions similaires (Liu, Henson, Keshavan, Pekka-Onnela et Torous, 2019), en plus de révéler la faisabilité d'évaluations longitudi-

nales facilitées par la ludification des tests neuropsychologiques, tels qu'implémentés par exemple dans l'application mobile LAMP (Torous et coll., 2019).

## **Données passives**

### **Activité physique, sédentarité et sommeil**

Comme mis de l'avant pour les Canadiens, l'activité physique régulière et modérée est associée à une meilleure santé mentale (Bernard et coll., 2018). L'inactivité physique est ainsi à la source d'une exacerbation des symptômes et d'une diminution de la qualité de vie dans la schizophrénie, la dépression ou l'anxiété (Hiles, Lamers, Milaneschi et Penninx, 2017). Plusieurs senseurs du téléphone intelligent (accéléromètre, gyroscope, GPS, état de l'écran, etc.) peuvent être utilisés avec succès pour prédire le type d'activité physique dans laquelle l'utilisateur est engagé (couché, assis, marche, course, montée/descente d'escalier, conduite automobile, etc.) (Vaizman, Ellis et Lanckriet, 2017). Les périodes de sommeil sont estimables à partir des mêmes senseurs, bien que l'architecture du sommeil ne puisse être adéquatement caractérisée sans avoir recours à la polysomnographie. Le suivi ambulatoire du sommeil à partir du téléphone intelligent dans les populations psychiatriques (Aledavood et coll., 2019) a une valeur clinique évidente au vu des perturbations du sommeil dans la plupart des troubles psychiatriques (Freeman, Sheaves, Waite, Harvey et Harrison, 2020).

### **Reconnaissance des émotions**

Des perturbations des affects, incluant l'humeur et les émotions, sont observées dans de très nombreux troubles psychiatriques (Bullis, Boettcher, Zavala, Farchione et Barlow, 2019). Grâce aux développements de l'informatique affective, le téléphone intelligent et ses senseurs ont le potentiel de faciliter la reconnaissance automatique des émotions pour plusieurs modalités d'expression (visage, voix, etc.) dans les troubles mentaux (Grabowski et coll., 2019).

Les expressions faciales, détectées par la caméra du téléphone, sont fréquemment caractérisées au travers du prisme de 6 émotions universelles: colère, peur, dégoût, joie, tristesse et surprise. Il est généralement considéré que les émotions sont transmises au travers de mouvements faciaux prototypiques qui peuvent être codés de façon

systématique (Martinez, Valstar, Jiang et Pantic, 2019). Différents algorithmes d'apprentissage automatique ont été entraînés à classer avec un certain succès le type d'émotion exprimée au travers des actions faciales, que ce soit à partir de photos statiques ou de séquences vidéo (Valstar, Mehu, Bihan Jiang, Pantic et Scherer, 2012). On notera cependant des défis propres à l'automatisation de la reconnaissance d'expressions faciales sur le téléphone intelligent, notamment en lien avec le mauvais cadrage ou une pauvre luminosité des images ou vidéos obtenues (Alshamsi, Meng et Li, 2016).

La voix, captée par le microphone du téléphone intelligent, est une autre modalité d'expression des émotions (Low, Bentley et Ghosh, 2020). De nombreuses mesures peuvent être extraites du signal vocal dans les domaines fréquentiel, d'énergie ou d'amplitude, spectral et temporel (Eyben et coll., 2016). L'efficacité d'un module de reconnaissance vocale et de classification des émotions sur le téléphone intelligent a par exemple été décrite pour le système *EmotionSense* (Rachuri et coll., 2010). Sur le plan de l'utilité clinique, l'analyse automatique du signal vocal obtenu du téléphone intelligent s'est avérée efficace pour la détection du stress (p. ex. *StressSense*, Lu et coll., 2012), des états dépressifs ou maniaques dans la dépression et le trouble bipolaire (p. ex. *Monarca*, Faurholt-Jepsen et coll., 2016), ainsi que d'autres troubles mentaux comme la schizophrénie où les symptômes négatifs se manifestent clairement au travers de l'analyse automatique de la voix (Low et coll., 2020).

Le contenu du langage écrit ou parlé, potentiellement extrait des messages texte, courriels ou médias sociaux dans un cas et des appels vocaux ou son ambiant dans l'autre, peut être soumis à des algorithmes relevant du traitement automatique du langage. En particulier, les analyses dites des sentiments visent à caractériser les états affectifs en y assignant une polarité (positive, négative, neutre) ou en classant les émotions exprimées (Messias et coll., 2017). À titre d'illustration de la pertinence clinique de cette modalité, l'analyse automatique du contenu linguistique de publications postées sur les médias sociaux a permis de détecter le niveau de stress de l'utilisateur (Pillai, Thelwall et Orasan, 2018), et de prédire un diagnostic ultérieur de dépression majeure (Eichstaedt et coll., 2018) ou un épisode psychotique à venir (Birnbaum et coll., 2019).

Les interactions tactiles avec l'écran du téléphone intelligent, par appuis ou glissements des doigts, représentent une approche alternative pour caractériser l'humeur et les émotions (p. ex. *TouchSense*,

Ghosh, Hiware, Ganguly, Mitra et De, 2019). La vitesse des appuis, sa variabilité, les pauses ou la pression des frappes sont autant de mesures, parmi d'autres, dont il a été montré qu'elles peuvent prédire le niveau de stress (Exposito, Hernandez et Picard, 2018) ainsi que la sévérité et les fluctuations des symptômes dans des troubles psychiatriques comme la dépression (Mastoras et coll., 2019) et le trouble bipolaire (p. ex. *Biaffect*, Zulueta et coll., 2018).

### Activité sociale

Les communications, appels et messages texte, sont une première fenêtre sur les comportements sociaux. La fréquence et la durée ou longueur des communications reçues et envoyées peuvent être extraites, au même titre que le nombre d'appels manqués ou nombre de contacts différents avec qui un échange est établi (Eskes, Spruit, Brinkkemper, Vorstman et Kas, 2016). Une quantification des interactions sociales est possible au travers de métriques de plus haut niveau, par exemple la diversité et la nouveauté des communications, la force du lien de communication ou encore la réciprocité de celles-ci (Bati et Singh, 2018). Les communications informent par exemple sur la solitude subjective (Pulekar et Agu, 2016) et le risque de rechute dans la schizophrénie (Buck et coll., 2019).

La technologie Bluetooth permet de capturer les interactions sociales en face à face (Sofia, Firdose, Lopes, Moreira et Mendes, 2016), en détectant d'autres appareils dont le Bluetooth est activé à l'intérieur d'un rayon de quelques mètres. La densité des interactions sociales, les relations d'amitié et la structure plus complexe de réseaux sociaux peuvent être ainsi identifiées (Eskes et coll., 2016). Les données GPS contribuent également à quantifier la proximité physique des membres d'une dyade ou d'un réseau. La géolocalisation apporte de surcroît des informations sur la diversité spatiale des interactions, la nature de lieux fréquentés (achalandage, public, etc.), et les comportements de mobilité de l'utilisateur du téléphone intelligent (Bati et Singh, 2018). Les mesures obtenues au travers des données Bluetooth et GPS sont fréquemment combinées à d'autres senseurs pour décrire l'activité sociale. En particulier, l'analyse automatique des données du microphone peut servir à détecter la présence de voix humaines et de conversations dans l'environnement de l'utilisateur. Ensemble, les différentes mesures de l'activité sociale obtenues au travers de ces senseurs se sont révélées utiles pour expliquer divers symptômes rencontrés chez des utilisateurs souffrant de troubles psychotiques, de l'humeur ou anxieux (Boukhechba et coll., 2017 ; Jongs et coll., 2020).

## Fonctions cognitives

Des mesures passives ont été décrites comme utiles pour faire des inférences sur le fonctionnement cognitif. Premièrement, les différentes formes d'interactions avec l'écran tactile du téléphone intelligent sont ensemble prédictives de la performance à de nombreux tests neuropsychologiques classiques (Dagum, 2018). Deuxièmement, le traitement automatique du langage, en plus d'être utile à la prédiction des émotions, permet d'évaluer plusieurs symptômes cognitifs dans les troubles psychiatriques. Les anormalités langagières sont au cœur des troubles de la pensée formelle, fortement prévalents dans la schizophrénie, mais également présents dans d'autres troubles mentaux (Kircher, Bröhl, Meier et Engelen, 2018). Le discours dépeint au travers de mesures dites de cohérence sémantique (p. ex. liées à la tangentialité ou au déraillement) et de complexité syntaxique (p. ex. révélatrices de la pauvreté du discours ou du contenu) a notamment été étudié avec succès dans la schizophrénie (Corcoran et coll., 2018).

## Perspectives

### Double rôle pour l'apprentissage automatique

Nombre des mesures de phénotypage digital que nous avons présentées résultent de la transformation par l'apprentissage automatique des données originellement massives et bruitées des senseurs du téléphone intelligent en des mesures intégratives du comportement, des émotions, de l'activité sociale et de la cognition (figure 1c). Bien que le phénotypage digital offre une nouvelle richesse à la mesure des troubles mentaux, les mégadonnées émanant ainsi du téléphone intelligent rendent d'autant plus complexe toute tâche prédictive pour le clinicien, du fait de leur nature hautement dimensionnelle et densément longitudinale. Prédire une rechute ou une aggravation de l'état mental d'un patient est une tâche notoirement difficile (Cohen et coll., 2019). Des modèles prédictifs, basés sur l'apprentissage automatique qui s'accommodent particulièrement bien de cette complexité dimensionnelle dans l'objectif de poser des prédictions justes et généralisables à de nouveaux individus, pourraient supporter le clinicien en santé mentale dans sa prise de décision clinique (Shortliffe et Sepúlveda, 2018). L'apprentissage automatique est donc utile à 2 niveaux distinguables : d'une part, pour extraire des mesures de phénotypage digital résumant l'information massive émise par les senseurs, d'autre part, pour élaborer des modèles

prédictifs d'états mentaux individuels futurs reposant sur des mesures de phénotypage digital intégratives (Dwyer, Falkai et Koutsouleris, 2018 ; Mohr et coll., 2017) (figure 1b, 1c).

Encore relativement peu nombreuses en raison de l'émergence récente du domaine, les démonstrations d'application des méthodes d'apprentissage automatique sur des données de phénotypage digital à des fins de prédictions individuelles d'états mentaux futurs n'en sont pas moins prometteuses (Busk et coll., 2020 ; Spathis, Servia-Rodriguez, Farrahi, Mascolo et Rentfrow, 2019 ; Umematsu, Sano et Picard, 2019 ; Wang et coll., 2020). Particulièrement performantes sont les applications de l'apprentissage profond, incluant les réseaux de neurones récurrents capables de capturer les phénomènes dynamiques de données de phénotypage digital densément longitudinales (Durstewitz, Koppe et Meyer-Lindenberg, 2019 ; Koppe, Guloksuz, Reininghaus et Durstewitz, 2019), qui surpassent les formes plus classiques d'apprentissage automatique lorsque des données suffisamment massives sont disponibles pour l'entraînement des modèles. En contrepartie, les modèles prédictifs bâtis avec l'apprentissage profond, qui repose sur de nombreuses représentations non linéaires au travers de multiples couches, souffrent d'un manque d'interprétabilité (explicabilité) résultant en des « boîtes noires » dont les raisons des prédictions sont opaques à leurs utilisateurs (Murdoch, Singh, Kumbier, East-Richard et Yu, 2019). Il est cependant à noter que tous les types d'algorithmes d'apprentissage automatique ne souffrent pas d'un manque de transparence au même degré, et que d'importants efforts sont actuellement investis à développer des méthodes permettant une meilleure interprétabilité, ou explicabilité, des modèles et de leurs prédictions (Stiglic et coll., 2020 ; Lundberg et Lee, 2017).

### **Barrières à l'adoption dans la pratique clinique**

L'actuelle opacité des modèles d'apprentissage automatique représente un premier frein évident à l'emploi d'outils prédictifs reposant sur le phénotypage digital dans la pratique clinique en santé mentale. Le règlement général sur la protection des données mis en œuvre par le gouvernement européen en 2018 définit le droit de la personne dont émanent les données à ne pas être affectée par une décision ayant pour seul fondement leur traitement automatique<sup>3</sup>, soulignant par

---

3. <https://gdpr-info.eu/art-22-gdpr/>

là même le droit à une explication de la prédiction. D'autres défis éthiques et légaux devront être relevés pour permettre une adoption à grande échelle des technologies en question (Martinez-Martin, Insel, Dagum, Greely et Cho, 2018; Wiens et coll., 2019). Pour que la balance entre les bénéfices et les risques perçus penche suffisamment en faveur des premiers, il faudra s'assurer de la régulation stricte de la technologie proposée des points de vue de la responsabilité d'efficacité et de sûreté des produits, de la protection des données en y facilitant l'accès et le contrôle à l'utilisateur, de la transparence concernant la collection des données en plus de celle liée aux modèles prédictifs qui en émergent, de l'atteinte à la vie privée et du consentement éclairé, notamment.

La capacité à satisfaire ces exigences impactera naturellement la propension des patients à adopter les technologies proposées. Dans 2 études distinctes, une majorité de patients souffrant de troubles anxieux et de l'humeur se disaient intéressés à utiliser une application mobile de phénotypage digital à visée clinique dans le champ de la santé mentale (Di Matteo, Fine, Fotinos, Rose et Katzman, 2018; Nicholas et coll., 2019). On notera que l'ouverture à partager les données captées par l'application mobile est plus marquée lorsque le destinataire est le clinicien traitant plutôt que la famille (Nicholas et coll., 2019). Alors que les patients se disent généralement confortables à utiliser des applications mobiles permettant d'autorapporter les symptômes, ils émettent sans surprise plus de réserves concernant la collecte des données passives des senseurs et informations d'exploitation du téléphone intelligent (Di Matteo et coll., 2018; Torous, Wisniewski, Liu et Keshavan, 2018). Les différentes formes de senseurs ne sont par ailleurs pas sur un pied d'égalité : un tiers des patients se disaient non ouverts à permettre l'enregistrement de données GPS ou des senseurs de mouvement alors qu'un peu plus de la moitié d'entre eux se disaient opposés à donner accès aux données audio ou au contenu des messages texte (Di Matteo et coll., 2018). Les patients sont ainsi plus confortables à partager des données liées à l'information sur la santé (activité physique, sommeil, émotion, etc.) que des données jugées plus personnelles (communication, activité sociale, etc.) (Nicholas et coll., 2019). L'inquiétude principale étant liée à la sécurité et la confidentialité des données, il peut être argumenté qu'une proportion plus importante de patients seraient enclins à utiliser des outils de phénotypage digital advenant la garantie que les données brutes les plus sensibles ne soient pas partagées en l'état, en intégrant la transformation des mesures brutes des senseurs

en des métriques de plus haut niveau sur le téléphone même (Di Matteo et coll., 2018).

L'avènement des nouvelles technologies confronte les cliniciens à un changement de culture professionnelle et refaçonne la relation patient-clinicien (Bourla et coll., 2018; Warraich, Califf et Krumholz, 2018). Comme pour toute nouvelle technologie majeure, certaines craintes et réticences se font sentir, en rythme avec le cycle sisyphéen des paniques technologiques (Orben, 2020). Tous les cliniciens ne se caractérisent pas par le même degré d'acceptabilité des outils en question. Tel qu'exploré dans un échantillon de 515 psychiatres majoritairement jeunes, un niveau d'acceptabilité moyen ne reflétait pas la variabilité importante observée entre différents profils de cliniciens (Bourla et coll., 2018). Aux côtés d'une moitié des répondants à majorité praticiens hospitaliers ou dans le privé rapportant une acceptabilité potentielle en fonction de la situation et du patient apparaissaient 2 autres types de profils opposés. Le premier, plutôt universitaire et d'orientation neurobiologique, présentait un niveau d'acceptabilité élevé alors que le deuxième, souvent résident en formation et d'orientation psychanalytique, rapportait un faible niveau d'acceptation. Les niveaux d'acceptabilité variaient selon la dimension s'y rattachant (utilité, exploitabilité, fiabilité, risque) et le type de mesure (actif, passif).

## **Conclusion**

Alors que la loi d'Amara, du futurologue américain du même nom, souligne notre tendance à surestimer l'incidence d'une nouvelle technologie à court terme, elle met aussi de l'avant notre propension à la sous-estimer à long terme. Il serait donc avisé de modérer nos attentes actuelles et d'anticiper de possibles désillusions dans un futur proche, pour mieux réfléchir aux bénéfices qui émergeront d'un plus lent processus de maturation du phénotypage digital couplé à l'apprentissage automatique dans le domaine de la santé mentale. Les promesses de cette nouvelle approche technologique pour une pratique clinique améliorée se réaliseront pleinement, de façon responsable sur les plans éthique, légal et sécuritaire, par l'implication de toutes les parties prenantes – patients, cliniciens, chercheurs, décisionnaires – dans son développement (Martinez-Martin et coll., 2018; Wiens et coll., 2019).

Ce travail est financé par une subvention chercheur-boursier Junior 1 du FRQS à Pierre Orban.

## RÉFÉRENCES

- Aledavood, T., Torous, J., Hoyos, A. M. T., Naslund, J. A., Onnela, J.-P. et Keshavan, M. (2019). Smartphone-Based Tracking of Sleep in Depression, Anxiety, and Psychotic Disorders, 1-9. <http://doi.org/10.1007/s11920-019-1043-y>
- Alshamsi, H., Meng, H. et Li, M. (2016). Real time facial expression recognition app development on mobile phones (pp. 1750-1755). *12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)*. <http://doi.org/10.1109/FSKD.2016.7603442>
- American Psychiatric Association. (2016). *The American Psychiatric Association Practice Guidelines for the Psychiatric Evaluation of Adults* (3<sup>rd</sup> ed.). American Psychiatric Publishing, Inc.
- Bati, G. F. et Singh, V. K. (2018). "Trust Us": Mobile Phone Use Patterns Can Predict Individual Trust Propensity (pp. 1-14). *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: ACM Press. <http://doi.org/10.1145/3173574.3173904>
- Baumeister, H. et Montag C. (2019). Digital phenotyping and mobile sensing. *New developments in psychoinformatics*. Springer.
- Bernard, P., Doré, I., Romain, A.-J., Hains-Monfette, G., Kingsbury, C. et Sabiston, C. (2018). Dose response association of objective physical activity with mental health in a representative national sample of adults: A cross-sectional study. *PLoS One*, 13(10), e0204682-16. <http://doi.org/10.1371/journal.pone.0204682>
- Birnbaum, M. L., Ernala, S. K., Rizvi, A. F., Arenare, E., Van Meter, A., Choudhury, M. et Kane, J. M. (2019). Detecting relapse in youth with psychotic disorders utilizing patient-generated and patient-contributed digital data from Facebook. *Npj Schizophrenia*, 1-9. <http://doi.org/10.1038/s41537-019-0085-9>
- Boukhechba, M., Huang, Y., Chow, P., Fua, K., Teachman, B. A. et Barnes, L. E. (2017). Monitoring social anxiety from mobility and communication patterns (pp. 749-753). *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, New York, NY, USA: ACM. <http://doi.org/10.1145/3123024.3125607>
- Bourla, A., Ferreri, F., Ogorzelec, L., Peretti, C.-S., Guinchard, C. et Mouchabac, S. (2018). Psychiatrists' Attitudes Toward Disruptive New Technologies: Mixed-Methods Study. *JMIR Mental Health*, 5(4), e10240-12. <http://doi.org/10.2196/10240>
- Buck, B., Scherer, E., Brian, R., Wang, R., Wang, W., Campbell, A. et coll. (2019). Relationships between smartphone social behavior and relapse in schizophrenia: A preliminary report. *Schizophrenia Research*, 208, 167-172. <http://doi.org/10.1016/j.schres.2019.03.014>
- Bullis, J. R., Boettcher, H., Zavala, S. S., Farchione, T. J. et Barlow, D. H. (2019). What is an emotional disorder? A transdiagnostic mechanistic definition with implications for assessment, treatment, and prevention. *Clinical Psychology: Science and Practice*, 26(2), e12278. <http://doi.org/10.1111/cpsp.12278>
- Busk, J., Faurholt-Jepsen, M., Frost, M., Bardram, J. E., Vedel Kessing, L. et Winther, O. (2020). Forecasting Mood in Bipolar Disorder From Smartphone

- Self-assessments: Hierarchical Bayesian Approach. *JMIR mHealth and uHealth*, 8(4), e15028-14. <http://doi.org/10.2196/15028>
- Chiu, M., Gatov, E., Vigod, S. N., Amartey, A., Saunders, N. R., Yao, Z. et coll. (2018). Temporal Trends in Mental Health Service Utilization across Outpatient and Acute Care Sectors: A Population-Based Study from 2006 to 2014. *The Canadian Journal of Psychiatry*, 63(2), 94-102. <http://doi.org/10.1177/0706743717748926>
- Cohen, A. S., Fedechko, T., Schwartz, E. K., Le, T. P., Foltz, P. W., Bernstein, J. et coll. (2019). Psychiatric Risk Assessment from the Clinician's Perspective: Lessons for the Future. *Community Mental Health Journal*, 55(7), 1165-1172. <http://doi.org/10.1007/s10597-019-00411-x>
- Corcoran, C. M., Carrillo, F., Fernández-Slezak, D., Bedi, G., Klim, C., Javitt, D. C. et coll. (2018). Prediction of psychosis across protocols and risk cohorts using automated language analysis. *World Psychiatry: Official Journal of the World Psychiatric Association (WPA)*, 17(1), 67-75. <http://doi.org/10.1002/wps.20491>
- Dagum, P. (2018). Digital biomarkers of cognitive function. *Npj Digital Medicine*, 1-3. <http://doi.org/10.1038/s41746-018-0018-4>
- Di Matteo, D., Fine, A., Fotinos, K., Rose, J. et Katzman, M. (2018). Patient Willingness to Consent to Mobile Phone Data Collection for Mental Health Apps: Structured Questionnaire. *JMIR Mental Health*, 5(3), e56. <http://doi.org/10.2196/mental.9539>
- Durstewitz, D., Koppe, G. et Meyer-Lindenberg, A. (2019). Deep neural networks in psychiatry. *Molecular Psychiatry*, 1-16. <http://doi.org/10.1038/s41380-019-0365-9>
- Dwyer, D. B., Falkai, P. et Koutsouleris, N. (2018). Machine Learning Approaches for Clinical Psychology and Psychiatry. *Annual Review of Clinical Psychology*, 14(1), 91-118. <http://doi.org/10.1146/annurev-clinpsy-032816-045037>
- East-Richard, C., R Mercier, A., Nadeau, D. et Cellar, C. (2020). Transdiagnostic Neurocognitive Deficits in Psychiatry: A Review of Meta-Analyses. *Canadian Psychology/Psychologie Canadienne*, 61(3), 190-214. <http://doi.org/10.1037/cap0000196>
- Eichstaedt, J. C., Smith, R. J., Merchant, R. M., Ungar, L. H., Crutchley, P., Preoțiu-Pietro, D. et coll. (2018). Facebook language predicts depression in medical records. *Proceedings of the National Academy of Sciences*, 115(44), 11203-11208. <http://doi.org/10.1073/pnas.1802331115>
- Eskes, P., Spruit, M., Brinkkemper, S., Vorstman, J. et Kas, M. J. (2016). The sociability score: App-based social profiling from a healthcare perspective. *Computers in Human Behavior*, 59(c), 39-48. <http://doi.org/10.1016/j.chb.2016.01.024>
- Exposito, M., Hernandez, J. et Picard, R. (2018). Affective keys: towards unobtrusive stress sensing of smartphone users (pp. 139-145). *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*. New York, NY, USA: ACM Press. <http://doi.org/10.1145/3236112.3236132>

- Eyben, F., Scherer, K. R., Schuller, B. W., Sundberg, J., Andre, E., Busso, C. et coll. (2016). The Geneva Minimalistic Acoustic Parameter Set (GeMAPS) for Voice Research and Affective Computing. *IEEE Transactions on Affective Computing*, 7(2), 190-202. <http://doi.org/10.1109/TAFFC.2015.2457417>
- Faurholt-Jepsen, M., Busk, J., Frost, M., Vinberg, M., Christensen, E. M., Winther, O. et coll. (2016). Voice analysis as an objective state marker in bipolar disorder. *Translational Psychiatry*, 6(7): e856. <http://doi.org/10.1038/tp.2016.123>
- Fortney, J. C., Ünützer, J., Wrenn, G., Pyne, J. M., Smith, G. R., Schoenbaum, M. et Harbin, H. T. (2017). A Tipping Point for Measurement-Based Care. *Psychiatric Services*, 68(2), 179-188. <http://doi.org/10.1176/appi.ps.201500439>
- Freeman, D., Sheaves, B., Waite, F., Harvey, A. G. et Harrison, P. J. (2020). Sleep disturbance and psychiatric disorders. *Lancet Psychiatry*, 7(7), 628-637. [http://doi.org/10.1016/S2215-0366\(20\)30136-X](http://doi.org/10.1016/S2215-0366(20)30136-X)
- Ghosh, S., Hiware, K., Ganguly, N., Mitra, B. et De, P. (2019). Emotion detection from touch interactions during text entry on smartphones. *Journal of Human Computer Studies*, 130, 47-57. <http://doi.org/10.1016/j.ijhcs.2019.04.005>
- Goldberg, S. B., Buck, B., Raphaely, S. et Fortney, J. C. (2018). Measuring Psychiatric Symptoms Remotely: a Systematic Review of Remote Measurement-Based Care. *Current Psychiatry Reports*, 20(10): 81 1-12. <http://doi.org/10.1007/s11920-018-0958-z>
- Grabowski, K., Rynkiewicz, A., Lassalle, A., Baron-Cohen, S., Schuller, B., Cummins, N. et coll. (2019). Emotional expression in psychiatric conditions: New technology for clinicians. *Psychiatry and Clinical Neurosciences*, 73(2), 50-62. <http://doi.org/10.1111/pcn.12799>
- Hiles, S. A., Lamers, F., Milaneschi, Y. et Penninx, B. W. J. H. (2017). Sit, step, sweat: longitudinal associations between physical activity patterns, anxiety and depression. *Psychological Medicine*, 47(8), 1466-1477. <http://doi.org/10.1017/S0033291716003548>
- Hirschtritt, M. E. et Insel, T. R. (2018). Digital Technologies in Psychiatry: Present and Future. *Focus*, 16(3), 251-258. <http://doi.org/10.1176/appi.focus.20180001>
- Jongs, N., Jagesar, R., Haren, N. E. M., Penninx, B. W. J. H., Reus, L., Visser, P. J. et coll. (2020). A framework for assessing neuropsychiatric phenotypes by using smartphone-based location data. *Translational Psychiatry*, 10, 211. <http://doi.org/10.1038/s41398-020-00893-4>
- Kircher, T., Bröhl, H., Meier, F. et Engelen, J. (2018). Formal thought disorders: from phenomenology to neurobiology. *Lancet Psychiatry*, 5(6), 515-526. [http://doi.org/10.1016/S2215-0366\(18\)30059-2](http://doi.org/10.1016/S2215-0366(18)30059-2)
- Koppe, G., Guloksuz, S., Reininghaus, U. et Durstewitz, D. (2019). Recurrent Neural Networks in Mobile Sampling and Intervention. *Schizophrenia Bulletin*, 45(2), 272-276. <http://doi.org/10.1093/schbul/sby171>
- Liu, G., Henson, P., Keshavan, M., Pekka-Onnela, J. et Torous, J. (2019). Assessing the potential of longitudinal smartphone based cognitive assessment in schizophrenia: A naturalistic pilot study. *Schizophrenia Research Cognition*, 17, 100144. <http://doi.org/10.1016/j.scog.2019.100144>

- Low, D. M., Bentley, K. H. et Ghosh, S. S. (2020). Automated assessment of psychiatric disorders using speech: A systematic review. *Laryngoscope Investigative Otolaryngology*, 5(1), 96-116. <http://doi.org/10.1002/lio2.354>
- Lu, H., Frauendorfer, D., Rabbi, M., Mast, M. S., Chittaranjan, G. T., Campbell, A. T. et coll. (2012). Stresssense: Detecting stress in unconstrained acoustic environments using smartphones (pp. 351-360). *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*. <http://doi.org/10.1145/2370216.2370270>
- Lundberg, S.M. et Lee, S.-I. (2017) A unified approach to interpreting model predictions (pp. 4768-4777). *Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS)*. <http://doi.org/10.5555/3295222.3295230>
- Martinez, B., Valstar, M. F., Jiang, B. et Pantic, M. (2019). Automatic Analysis of Facial Actions: A Survey. *IEEE Transactions on Affective Computing*, 10(3), 325-347. <http://doi.org/10.1109/TAFFC.2017.2731763>
- Martinez-Martin, N., Insel, T. R., Dagum, P., Greely, H. T. et Cho, M. K. (2018). Data mining for health: staking out the ethical territory of digital phenotyping. *Npj Digital Medicine*, 1-5. <http://doi.org/10.1038/s41746-018-0075-8>
- Mastoras, R.-E., Iakovakis, D., Hadjidimitriou, S., Charisis, V., Kassie, S., Alsaadi, T. et coll. (2019). Touchscreen typing pattern analysis for remote detection of the depressive tendency. *Scientific Reports*, 9(1), 13 414. <http://doi.org/10.1038/s41598-019-50002-9>
- Messias, J., Diniz, J. P., Soares, E., Ferreira, M., Araújo, M., Bastos, L. et coll. (2017). An evaluation of sentiment analysis for mobile devices. *Social Network Analysis and Mining*, 7: 20. <http://doi.org/10.1007/s13278-017-0437-2>
- Mohr, D. C., Zhang, M. et Schueller, S. M. (2017). Personal Sensing: Understanding Mental Health Using Ubiquitous Sensors and Machine Learning. *Annual Review of Clinical Psychology*, 13(1), 23-47. <http://doi.org/10.1146/annurev-clinpsy-032816-044949>
- Murdoch, W. J., Singh, C., Kumbier, K., East-Richard, C. et Yu, B. (2019). Definitions, methods, and applications in interpretable machine learning. *Proceedings of the National Academy of Sciences*, 116(44), 22071-22080. <http://doi.org/10.1073/pnas.1900654116>
- Myin-Germeys, I., Kasanova, Z., Vaessen, T., Vachon, H., Kirtley, O., Viechtbauer, W. et Reininghaus, U. (2018). Experience sampling methodology in mental health research: new insights and technical developments. *World Psychiatry*, 17(2), 123-132. <http://doi.org/10.1002/wps.20513>
- Nicholas, J., Shilton, K., Schueller, S. M., Gray, E. L., Kwasny, M. J. et Mohr, D. C. (2019). The Role of Data Type and Recipient in Individuals' Perspectives on Sharing Passively Collected Smartphone Data for Mental Health: Cross-Sectional Questionnaire Study. *JMIR mHealth and uHealth*, 7(4), e12578-10. <http://doi.org/10.2196/12578>
- Onnela, J.-P. (2021). Opportunities and challenges in the collection and analysis of digital phenotyping data. *Neuropsychopharmacology*, 46, 45-54. <http://doi.org/10.1038/s41386-020-0771-3>

- Orben, A. (2020). The Sisyphean Cycle of Technology Panics. *Perspectives on Psychological Science*, 15(5), 1143-1157. <http://doi.org/10.1177/1745691620919372>
- Passell, E., Dillon, D. G., Baker, J. T., Vogel, S. C. et Scheuer, L. S. (2019). Digital Cognitive Assessment: Results from the TestMyBrain NIMH Research Domain Criteria (RDoC) Field Test Battery Report. <https://doi.org/10.31234/osf.io/dcszr>
- Pillai, R. G., Thelwall, M. et Orasan, C. (2018). Detection of Stress and Relaxation Magnitudes for Tweets (pp. 1677-1684). *Companion proceedings of the The Web Conference*. <http://doi.org/10.1145/3184558.3191627>
- Pulekar, G. et Agu, E. (2016). Autonomously sensing loneliness and its interactions with personality traits using smartphones (pp. 134-137). *2016 IEEE Healthcare Innovation Point-Of-Care Technologies Conference (HI-POCT)*. <http://doi.org/10.1109/HIC.2016.7797715>
- Rachuri, K. K., Musolesi, M., Mascolo, C., Rentfrow, P. J., Longworth, C. et Aucinas, A. (2010). EmotionSense: a mobile phones based adaptive platform for experimental social psychology research (pp. 281-290). *Proceedings of the 12<sup>th</sup> ACM International Conference on Ubiquitous Computing*. <http://doi.org/10.1145/1864349.1864393>
- Rogler, L. H., Mroczek, D. K., Fellows, M. et Loftus, S. T. (2001). The neglect of response bias in mental health research. *The Journal of Nervous and Mental Disease*, 189(3), 182-187. <http://doi.org/10.1097/00005053-200103000-00007>
- Shortliffe, E. H. et Sepúlveda, M. J. (2018). Clinical Decision Support in the Era of Artificial Intelligence. *Jama*, 320(21), 2199-2. <http://doi.org/10.1001/jama.2018.17163>
- Sliwinski, M. J., Mogle, J. A., Hyun, J., Munoz, E., Smyth, J. M. et Lipton, R. B. (2018). Reliability and Validity of Ambulatory Cognitive Assessments. *Assessment*, 25(1), 14-30. <http://doi.org/10.1177/1073191116643164>
- Sofia, R., Firdose, S., Lopes, L. A., Moreira, W. et Mendes, P. (2016). NSense: A people-centric, non-intrusive opportunistic sensing tool for contextualizing nearness (pp. 1-6). *2016 IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom)*. <http://doi.org/10.1109/HealthCom.2016.7749490>
- Spathis, D., Servia-Rodriguez, S., Farrahi, K., Mascolo, C. et Rentfrow, J. (2019). Sequence Multi-task Learning to Forecast Mental Wellbeing from Sparse Self-reported Data (pp. 2886-2894). *Proceedings of the 25th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. New York, NY, USA: ACM. <http://doi.org/10.1145/3292500.3330730>
- Stiglic, G., Kocbek, P., Fijacko, N., Zitnik, M., Verbert, K. et Cilar, L. (2020). Interpretability of machine learning-based prediction models in healthcare. *WIREs Data Mining and Knowledge Discovery*, 10(5), e1379. <http://doi.org/10.1002/widm.1379>
- Torous, J., Wisniewski, H., Bird, B., Carpenter, E., David, G., Elejalde, E. et coll. (2019). Creating a Digital Health Smartphone App and Digital Phenotyping Platform for Mental Health and Diverse Healthcare Needs: an

- Interdisciplinary and Collaborative Approach. *Journal of Technology in Behavioral Science*, 4, 73-85. <http://doi.org/10.1007/s41347-019-00095-w>
- Torous, J., Wisniewski, H., Liu, G. et Keshavan, M. (2018). Mental Health Mobile Phone App Usage, Concerns, and Benefits Among Psychiatric Outpatients: Comparative Survey Study. *JMIR Mental Health*, 5(4), e11715. <http://doi.org/10.2196/11715>
- Umematsu, T., Sano, A. et Picard, R. W. (2019). Daytime Data and LSTM can Forecast Tomorrow's Stress, Health, and Happiness (pp. 2186-2190). *The 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. <https://doi.org/10.1109/EMBC.2019.8856862>
- Vaizman, Y., Ellis, K. et Lanckriet, G. (2017). Recognizing detailed human context in the wild from smartphones and smartwatches. *IEEE Pervasive Computing*, 16(4), 62-74. <http://doi.org/10.1109/MPRV.2017.3971131>
- Valstar, M. F., Mehu, M., Bihan Jiang, Pantic, M. et Scherer, K. (2012). Meta-Analysis of the First Facial Expression Recognition Challenge. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 42(4), 966-979. <http://doi.org/10.1109/TSMCB.2012.2200675>
- Wang, R., Wang, W., Obuchi, M., Scherer, E., Brian, R., Ben-Zeev, D. et coll. (2020). On Predicting Relapse in Schizophrenia using Mobile Sensing in a Randomized Control Trial (pp. 1-8). Proceedings of the 2020 IEEE International Conference on Pervasive Computing and Communications (PerCom). <http://doi.org/10.1109/PerCom45495.2020.9127365>
- Warraich, H. J., Califf, R. M. et Krumholz, H. M. (2018). The digital transformation of medicine can revitalize the patient-clinician relationship. *Npj Digital Medicine*, 1-3. <http://doi.org/10.1038/s41746-018-0060-2>
- Weizenbaum, E., Torous, J. et Fulford, D. (2020). Cognition in Context: Understanding the Everyday Predictors of Cognitive Performance in a New Era of Measurement. *JMIR mHealth and uHealth*, 8(7), e14328. <http://doi.org/10.2196/14328>
- Wichers, M. (2013). The dynamic nature of depression: a new micro-level perspective of mental disorder that meets current challenges. *Psychological Medicine*, 44(07), 1349-1360. <http://doi.org/10.1017/S0033291713001979>
- Wichers, M., Schreuder, M. J., Goekoop, R. et Groen, R. N. (2019). Can we predict the direction of sudden shifts in symptoms? Transdiagnostic implications from a complex systems perspective on psychopathology. *Psychological Medicine*, 49(3), 380-387. <http://doi.org/10.1017/S0033291718002064>
- Wiens, J., Saria, S., Sendak, M., Ghassemi, M., Liu, V. X., Doshi-Velez, F. et coll. (2019). Do no harm: a roadmap for responsible machine learning for health care. *Nature Medicine*, 25, 1337-1340. <http://doi.org/10.1038/s41591-019-0548-6>
- Wright, A. G. C. et Woods, W. C. (2020). Personalized Models of Psychopathology. *Annual Review of Clinical Psychology*, 16(1), 49-74. <http://doi.org/10.1146/annurev-clinpsy-102419-125032>

Zulueta, J., Piscitello, A., Rasic, M., Easter, R., Babu, P., Langenecker, S. A. et coll. (2018). Predicting Mood Disturbance Severity with Mobile Phone Keystroke Metadata: A BiAffect Digital Phenotyping Study. *Journal of Medical Internet Research*, 20(7), e241. <http://doi.org/10.2196/jmir.9775>