Tracking online poker problem gamblers with player account-based gambling data only

AMANDINE LUQUIENS,[1,2,3] MARIE-LAURE TANGUY,[4] AMINE BENYAMINA,[1,2,3] MARTHYLLE LAGADEC,[1,2,3] HENRI-JEAN AUBIN[1,2,3] & MICHEL REYNAUD[1,2,3]

1 Hôpital Paul Brousse, AP-HP, Villejuif, France
2 University Paris Sud, Paris, France
3 Inserm U 1178, Paris, France
4 URC, Hôpital Pitié –Salpetrière, AP-HP, Paris, France

Correspondence
Luquiens Amandine, Hôpital Paul Brousse, Addictologie, 12 avenue Paul Vaillant-Couturier, 94804 Villejuif Cedex, France.
Telephone (+33) 145 59 4018/3087
Fax (+33) 145 59 3863
Email: amandineluquiens@yahoo.fr

Abstract

The aim was to develop and validate an instrument to track online problem poker gamblers with player account-based gambling data (PABGD). We emailed an invitation to all active poker gamblers on the online gambling service provider Winamax. The 14,261 participants completed the Problem Gambling Severity Index (PGSI). PGSI served as a gold standard to track problem gamblers (i.e., PGSI ≥ 5). We used a stepwise logistic regression to build a predictive model of problem gambling with PABGD, and validated it. Of the sample 18% was composed of online poker problem gamblers. The risk factors of problem gambling included in the predictive model were being male, compulsive, younger than 28 years, making a total deposit > 0 euros, having a mean loss per gambling session > 1.7 euros, losing a total of > 45 euros in the last 30 days, having a total stake > 298 euros, having > 60 gambling sessions in the last 30 days, and multi-tabling. The tracking instrument had a sensitivity of 80%, and a specificity of 50%. The quality of the instrument was good. This study illustrates the feasibility of a method to develop and validate instruments to track online problem gamblers with PABGD only. *Copyright © 2016 John Wiley & Sons, Ltd.*

Introduction

Online problem gambling presents an increasing challenge to health care providers due to its significantly increasing prevalence (Costes *et al.,* 2011; Wardle *et al.,* 2010) and the poor treatment-seeking behavior of affected subjects (Gainsbury *et al.,* 2014). Online gambling may be more likely than offline environments to contribute to problem gambling (Griffiths *et al.,* 2009). Applying responsible gambling principles to online gambling is still an unsolved

challenge, in terms of "who's in charge of", "who to track for" and "when to do it". Gambling disorders, as other addictions, are characterized by loss of control of the behavior (Piazza and Deroche-Gamonet, 2013; Tang *et al.,* 2015). This loss of control can be illustrated through different symptoms and at different level of severity. Several other terms have been previously proposed to designate these levels, such as at-risk gamblers, problem gamblers, or pathological gamblers. In this article, we chose to focus our interest on problem gambling in an inclusive meaning (i.e. subclinical problem gambling, problem gambling and pathological gambling), reflecting a continuum characterized by the progressively increase of behavior loss of control.

Predictive factors of problem online poker gambling have been described, including time devoted to gambling and psychological states (Hopley and Nicki, 2010; Barrault and Varescon, 2013). An increasing body of literature has proposed cues to profile problem online gamblers based on their gambling behavior. Several authors have proposed identifying problem gamblers through routine data registered by online gambling service providers, using player account-based gambling data (PABGD) (Gainsbury, 2011; Griffiths and Whitty, 2010). Many papers illustrates this possible academic-provider partnership, for instance through the "Transparency Project", analyzing the account-player dataset of "bwin" provider: Braverman *et al.* (2013) described four categories of predictive factors of problem gambling: summary variables, dynamic variables, calendar-related variables and data related to gambling behavior before the first deposit, also called pre-deposit data. Summary variables were defined as descriptive variables that summarize gambling activities: the total active days, the sum of stakes/bets, the mean of stakes/bets (LaBrie and Schaffer, 2011), the variety of different stakes, the number of various games played, the variety of different games, an increase in time spent gambling, frequent payment method changes, verbal aggression and complaints to customer service (Griffiths and Whitty, 2010), and exceeded deposit limits (Broda *et al.,* 2008). Dynamic factors were defined as a change in gambling behavior patterns, which could express a loss of control, a gradual increase in stakes over time, repeated chasing patterns, and intense activity following subscription to an online gambling service (LaBrie and Schaffer, 2011). Despite these factors, the systematic tracking of at-risk gamblers has remained difficult due to ethical and methodological issues (Griffiths and Whitty, 2010). Dynamic variables, pre-deposit data and calendar-related variables may be poorly applied generally and may be difficult to include in repeated systematic screening of all

gamblers due to feasibility issues. Moreover, the methodology used to identify these factors has limitations. In all studies but one, the authors used proxies of gambling disorder but did not use a clinical instrument to differentiate problem and pathological gamblers from non-problem gamblers (LaPlante *et al.,* 2014). Previously used proxies have included alerts from providers' responsible gambling programs, which often involve highly subjective criteria and are neither systematically nor automatically useable, such as repeated customer complaints about fair play in an extremely agitated manner, contact from a relative of the customer to the provider to have the account blocked due to irresponsible gambling, a user's request for a scheduled payout to be cancelled after the user issued a payout in the portal interface (Braverman *et al.,* 2013), and a user's tentativeness to exceed gambling moderators (Broda *et al.,* 2008). Other previously used proxies have involved substantial involvement with problem and pathological gambling. Involvement is defined as financial involvement (e.g. high stakes or losses) (LaPlante *et al.,* 2009), temporal depth involvement (e.g. number of gambling days) or breadth of involvement in multiple games. The cutoff often relies on the distribution of these variables in the population of included gamblers, which could be a limitation due to recruitment bias and the operational choice of this cutoff (e.g. 5% of the largest total wagered). Temporal depth involvement has recently been shown to be insufficient as a unique criterion to predict gambling disorders (LaPlante *et al.,* 2014). More generally, multiple indicators are usually required to reliably identify problem and pathological gamblers (Delfabbro *et al.,* 2012).

Most operators propose virtual harm reduction efforts, such as self-limited gambling expenditure, and some of them have had effects on intensive gamblers (Auer and Griffiths, 2013). Despite these initiatives, guidelines for responsible gambling standards (National Council on Problem Gambling, 2012) and legal requirements of proposing moderators for gambling providers when opening an account in some countries (Assemblée Nationale, 2010), the prevalence of problem gambling has not decreased. Some applications as Playscan (Griffiths *et al.,* 2009) have been proposed to screen at-risk gamblers through online gambling behavior. However, to date, no tool, including Playscan, has been built comparing PABGD to a clinical screening tool. Thus, our tool completes the work initiated with previous tools using previously described literature or hypothetical factors. Moreover, Playscan and other tools require a personal initiative to screen one's gambling behavior. However, if systematic targeted interventions among the problem gamblers could enhance the efficacy of existing measures, their effect would

remain limited due to the difficulty identifying online problem gamblers and the need with existing screening methods of early personal initiative to be screened, when most of gamblers starting their gambling practice do not feel themselves as vulnerable for problem gambling.

This study is the first part of an online randomized controlled trial among problem gamblers with naturalistic recruitment; most trials of behavioral interventions have recruited patients from advertisements or from health care centers (Fink *et al.,* 2012). It was conducted to address the need to track and enhance access to care for online problem gamblers. It relies on an academic–industrial partnership. Its aim is to illustrate the feasibility of a method which proposes to develop and validate a sensitive and easy-to-use instrument to track poker-playing problem gamblers in the gambling environment by systematically analyzing their gambling behavior with no personal initiative required. Such an instrument would allow an automatic and systematic screening independent from gambling service providers, without a particular initiative of the gambler, to detect problem gambling, using the PABGD securely saved by the French authorities routinely. It is the first report of the development of a tool by confronting player account-based behavioral gambling data and a clinical screening self-questionnaire, the Problem Gambling Severity Index (PGSI) (Ferris and Wynne, 2001). It involves two successive objectives: (1) a comparison of non-problem and problem gamblers' characteristics at baseline; (2) the development of a specific and sensitive screening instrument for the tracking of problem gambling by modeling the correlation between routine provider-registered data and a widespread clinical diagnosis instrument, the PGSI, which is part of the Canadian Problem Gambling Index.

Methods

This study was divided into two overlapping phases. The first was a cross-sectional phase that screened all active poker gamblers on the poker gambling service provider Winamax. The second and interventional phase was an exclusively online randomized controlled trial that included the poker-playing problem gamblers identified in the first phase. In our study we chose to use the term "problem gambler" in an inclusive meaning, designating moderate risk (i.e. subclinical forms of problem gambling) and problem gamblers, including pathological gamblers, as mentioned by the National Center for Responsible Gaming (Reilly and Smith, 2013). The methods and results of the interventional phase will be presented in a separate publication.

Subject recruitment

Subjects were systematically proposed to be included in the cross-sectional phase of the study when they started a poker session during the inclusion period, November 13, 2013 to January 16, 2014 (subjects could be included only once). The other inclusion criteria were age ≥ 18, completed registration (i.e. an identification card was sent to Winamax to confirm age), and registration for ≥ 30 days. The day after players first opened a poker session during the inclusion period, they were sent an email that contained a link through which they were invited to connect to an online survey platform hosted by Winamax, where data were collected and then provided to an investigator. Other routinely registered gambling data were collected separately by Winamax and provided to the investigator. Winamax was contractually commissioned to collect the data, but the authors analyzed the data independently from the operator Winamax.

The assessment of enrolled subjects was completed exclusively online. Subjects' consent was obtained as required by local French laws and regulations. The study was authorized by the "Comité de Protection des Personnes" as required for medical interventional research in France. The subjects did not receive any compensation for their participation in this study. Subject anonymity was established and maintained throughout the course of the study. Before completing the online process, the subjects read a page that contained clear information about the phase in which they were to be included. The subjects had to read the page to confirm that they agreed and that they understood the study to proceed to the survey.

Sample size

The sample size of the cross-sectional phase was derived from the screening needs of the interventional phase. The sample size for the interventional phase was 992 patients. We systematically recruited gamblers to be included in the study until we attained the desired sample size for the interventional phase (Figure 1). We then included 14,261 gamblers, and the sample was divided into a learning group ($n = 9998$) to build the model and a testing group ($n = 4263$) to validate the model.

Settings and data collection

After gamblers opened a gambling session on the Winamax website, they were automatically emailed an invitation to complete the study. The baseline data that were prospectively collected automatically by the operator Winamax were retrospectively extracted for the 30-day
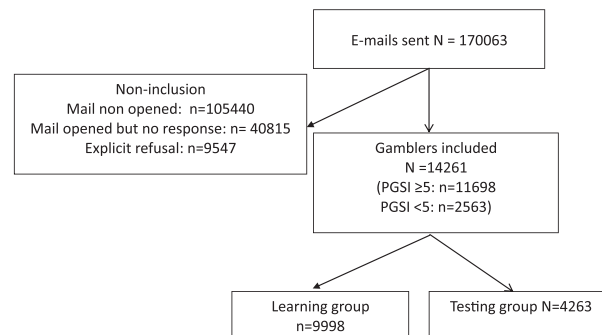
Figure 1. Flow chart.

period before the inclusion day. The authors conducted the data management and analysis.

Measures

The only additional data collected online involved the PGSI (Ferris and Wynne, 2001; Miller *et al.,* 2013). Since its publication in 2001, the PGSI has become internationally recognized as a robust measure of gambling behavior and has been used in Canada, Australia, Great Britain, Iceland and Norway. A new screening cutoff for this index has recently been proposed (Currie *et al.,* 2010). Results show that two of the PGSIs subtypes – the non-problem and problem gambling subtypes – represent distinct and well-defined groups, but that the original classification of low-risk (score 1–3) and moderate-risk (score 4–7) subtypes fits less well. Currie *et al.* (2010) and recommended eliminating the low-risk and moderate-risk subtypes in favor of two new mid-level categories consisting of low-risk gamblers, defined as a PGSI score of 1 to 4, and moderate gamblers, defined as a PGSI score of 5 to 7. Gamblers with a score of 3 and 4 were a loss homogeneous group. Others are suggesting that the PGSI cut-point of 8 is too stringent and recommend 5 and above to define problem gamblers (Williams *et al.,* 2012). We chose *a priori* to use this new threshold of 5 for problem gambling, as encouraged by the Canadian Consortium for Gambling Research (2013), to be more conservative than the previous threshold of 3 for at risk of problem gambling and to avoid inadequate sensibility of the instrument. Moreover, the choice of a conservative threshold was driven by the nested clinical trial proposing a therapeutic intervention to the identified problem gamblers.

Basic socio-demographic and routine registered data were extracted from the Winamax player account-based dataset. The gambling criteria used were previously reported to be good indicators of problem and pathological gambling (Griffiths and Whitty, 2010), and this information is routinely registered by the provider Winamax. We selected the gambling variables on the ease of their extraction from the routine database according to Winamax and their reproducibility among the other online gambling providers. Our goal was to build a model that could be routinely used for screening purposes, including variables required by French regulators for online gambling providers. We chose a short recall period of 30 days to rapidly implement a therapeutic intervention after the identification of problem gambling. These criteria were multi-tabling in the last 30 days (yes/no), compulsivity (yes/no) (defined by at least three deposits in a period of 12 hours), amount of total deposit in the last 30 days (euros) (an initial deposit is required upon opening the gambling account, which implies that some gamblers could have a null deposit during the study period), mean of loss per gambling session including the rake (euros), total loss in the last 30 days including the rake (euros), total stakes (euros), number of gambling sessions in the last 30 days, number of gambling days in the last 30 days, and time gambled (hours) in the last 30 days. Two criteria could not be correctly assessed for 57 gamblers due to technical limits linked to a lack of automatic disconnection from the application on some wireless devices, especially smartphones and tablet computers: time gambled (hours) in the last 30 days and number of gambling days in the last 30 days. We decided to exclude these two criteria from the predictive model because no technical solution could be identified to avoid this issue, which meant that the providers were unable to correctly collect information on these two criteria.

Statistical analysis

Descriptive analyses were performed for the total population and by group. The groups were (A) non-problem gamblers (i.e. PGSI < 5) and (B) problem gamblers (i.e. PGSI ≥ 5). Quantitative variables were compared using Student's *t*-test,

and qualitative variables were compared with the chi-squared test. The Hochberg method was used to address the multiple testing problem.

The sample was randomly divided into two groups. The first group (the learning group) was composed of 70% of the subjects and was used to build the model that explained problem gambling, as defined by PGSI ≥ 5. To identify the predictive factors of problem gambling (PGSI score ≥ 5), we used a stepwise logistic regression model. The factors introduced in the model were sex, age and all gambling criteria mentioned earlier. The quantitative variables were categorized into quartiles. The second group (the testing group) was composed of 30% of the sample and was used to validate the model. We used a split-sample approach to assess the performance of the prediction model because of the very large number of subjects included in the study. The goodness-of-fit of the model was assessed for the learning and validation groups with the Hosmer–Lemeshow test (Figure 2) to evaluate the calibration and the area under the receiver operating characteristic (ROC) curve (AUC) as a measure of discrimination.

Using the model's regression equation, we computed the predictive probability that a subject would be a problem gambler. We selected a probability value as a cutoff point to determine whether a subject was considered a problem gambler to obtain a sensitivity of 80% to limit the number of false negative subjects. To assess the performance of the decision rule with this cutoff point, a classification table was generated.

The included gamblers and non-included gamblers were compared using Student's *t*-tests and chi-squared tests.

All tests were two-sided. The alpha level was set at 0.05.

Results

The study was systematically proposed by email to 170,063 poker gamblers. We included 14,261 poker gamblers in the study. Non-inclusion criteria are detailed in the flow chart in Figure 1. Non-included gamblers' gambling data were collected and presented very close means on all criteria for the included population.

We found that 18% of the participants were problem gamblers, as defined by PGSI ≥ 5 ($n = 2563$). In comparison with previous literature, we found that 38% of gamblers had a PGSI score ≥ 3. As expected, the rate of problem gamblers was significantly lower among the female population (13.4% versus 18.4%, p-value = 0.0001). The gamblers' characteristics in the entire population and by group (PGSI < or ≥ 5) are detailed and compared in Table 1. The mean amount of money that poker gamblers lost was small, less than 1 euro per gambling session, and they lost less than 25 euros in the month. However, we found a significant difference between the two groups on every criterion, including the financial criteria. On average, non-problem gamblers did not lose money in a gambling session, whereas problem gamblers lost 3 euros on average.

The risk factors for problem gambling are being male (versus female), being younger than 28, making a total deposit > 0, having a mean loss per gambling session > 1.7, having a total loss in the last 30 days > 45 euros, having a total stake > 298 euros, playing > 60 gambling sessions in the last 30 days, multi-tabling and being compulsive. These combined factors constituted the predictive model of problem poker gambling (Table 2).

We tested the quality of this model on the learning group. The Hosmer–Lemeshow test indicated good calibration ($p = 0.52$). The discrimination assessed by the AUC was good (0.72). The model was then validated in the second remaining sample ($n = 4263$). The Hosmer–Lemeshow test was statistically significant, but the AUC was still good, 0.73 (Figure 3). The misclassification rate with a threshold that maximized the sum of sensitivity
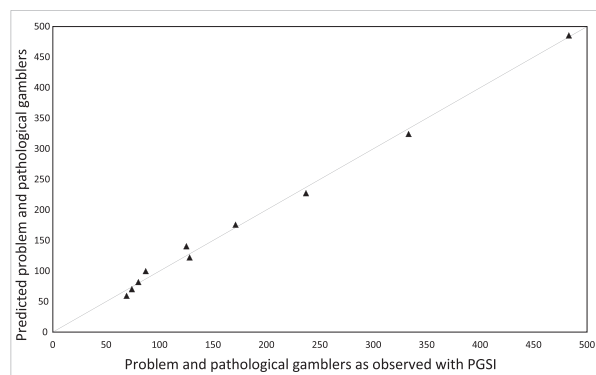
Figure 2. Hosmer–Lemeshow test (p-value = 0.52) (non-significance is expected from a valid model) ($n = 9998$).

Table 1. Gamblers characteristics and comparison between the two groups (PGSI < 5 and PGSI ≥ 5)

|	All N = 14261	PGSI < 5 n = 11698	PGSI ≥ 5 n = 2563
Age (years)*	35.8	36 (11)	33.5 (9)
Gender (men)*	90%	89.5%	92.5%
Total deposit in the last 30 days (mean, SD) (euros)*	95.2 (453)	61.7 (362.2)	248.1 (717.4)
Loss per gambling session including the rake (mean, SD) (euros)*	0.7 (23.3)	0.1 (24.7)	3.2 (15.1)
Total loss in the last 30 days including the rake (mean, SD) (euros)*	24.9 (585.5)	3.5 (556.7)	122.5 (694.1)
Total stake (mean, SD) (euros)*	814.78 (6335.2)	631.2 (5374.4)	1652.9 (9521.4)
Number of gambling days in the last 30 days (mean, SD)*	13.2(9.8)	12.8 (9.7)	14.8 (9.9)
Number of gambling sessions in the last 30 days (mean, SD)*	45.3 (58.5)	42.1 (53.5)	60.1 (75.6)
Time gambled in the last 30 days (mean, SD) (hour)*	49.9 (62.5)	47.1 (60.2)	62.8 (70.4)
Multitabling in the last 30 days (yes) *	61.2%	57.4%	78.2%
Compulsivity (yes)*	2.2%	1.6%	7.4%
PGSI score (mean, SD)*	2.6 (3.4)	1.4 (1.3)	8.3 (4.3)

*Adjusted *p*-value < 0.0001 (Hochberg method).

Table 2. Multivariate analysis explaining PGSI ≥ 5 (logistic regression) (learning group, n = 9998)

	Estimate	*p*-Value
Age ≥ 34 years old	−0.21	<0.001
Age ≥ 28 years old	−0.49	<0.0001
Women versus men	−0.24	0.02
Total deposit in the last 30 days ≥ 60 euros	0.33	<0.001
Total deposit in the last 30 days ≥ 0 euros	0.46	<0.0001
Mean of theoretical loss per gambling session including the rake ≥ 1.7 euros	0.39	<0.0001
Theoretical loss in the last 30 days including the rake ≥ 45 euros	0.23	0.01
Total stake ≥ 298 euros	0.46	<0.0001
Number of gambling sessions in the last 30 days ≥ 60	0.19	<0.01
Multitabling (yes)	0.32	<0.0001
Compulsivity (yes)	0.57	<0.0001

and specificity was 31.5%. We chose a sensitivity of 80% because the aim of the model was to be a screening test. The threshold was thus chosen to be 0.092, and the specificity decreased to 49.3%. The classification table at this threshold is shown in Table 3. The correctly classified gambler rate was 55%, the negative predictive value was 91.9%, and the positive predictive value was 26.1%.

We further explored the misclassified gamblers (i.e. the false positives; 1767, 50.7%). Among the false-positive gamblers, 31% had a Canadian PSGI score ≥ 3, which is usually considered problem gambling (moderate risk). More than 75% of false positives responded positively to at least one question on the PGSI. This means that only 10.1% of all gamblers were predicted to be problem
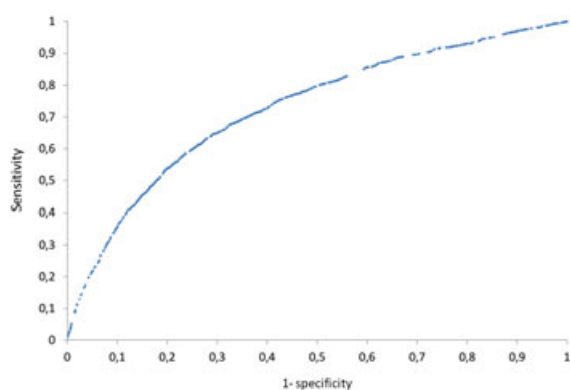
Figure 3. ROC curve on the validation sample (n = 4263).

Table 3. Classification table (testing group, $n = 4263$)

	Predicted		
	Problem gamblers	Non-problem gamblers	Total
Actual			
PGSI ≥ 5	*624*	152	776
PGSI < 5	1767	*1720*	3487
Total	2391	1872	4263

Note: Italic typeface corresponds to correctly classified gamblers.

gamblers by the model but were considered "not at risk" based on the classical PGSI score interpretation.

Discussion

Due to the large enrollment, we illustrated the feasibility of our method by developing and validating a model predictive of the PGSI status with good sensitivity of 80% and an acceptable rate of 10% of not-at-risk gamblers misclassified as poker problem gamblers, among active poker gamblers on the Winamax site.

Our study presents several strengths. It is the first report of the development of a tool by confronting player account based behavioral gambling data and a clinical screening self-questionnaire. It allowed the identification of player-account-based gambling variables predicting the problem gambling status as assessed by a clinical tool, using all available and easily extractable variables that were significantly and independently associated to problem gambling. Our results support previously described summary variables as predictive factors of problem gambling, particularly temporal involvement, here approximated through number of sessions, the sum of stakes, the mean of stakes (LaBrie and Schaffer, 2011), and breadth involvement, measured by a proxy within poker practice as multi-tabling (LaPlante *et al.*, 2014). Particularly, the total loss of 122.5 euros in 30 days corresponds to the level of loss previously described by LaPlante *et al.* (2014) in a longer period of two years ($122.5 \times 24 = 2940$ versus 2888 euros). This finding offers a strong argument to support our choice of a recall period of 30 days, which seems sufficient to obtain relevant information and short enough to react in its *in situ* use. In addition to summary variables, our model included a dynamic variable, compulsivity, which can be considered a behavioral expression of a chasing pattern. Our study identified several cutoffs in gambling variables that could be useful in clinical practice and prevention. Particularly, it lead to the identification of very

low thresholds in term of financial involvement, which is also a critical finding, at least in the poker gambling field. Despite the widespread image of problem gambling, our study showed that people with problem gambling can suffer from few smaller financial consequences. This is an important paradigm that may be specific to poker problem gambling and that should be taken into account for future therapeutic and preventive interventions. Our results support our concept of problem gambling as a combination of symptoms, with different levels of severity, illustrating loss of control of the behavior. They support the previously initiated studies proposing complex predictive model for problem gambling relying on gambling data (Griffiths *et al.,* 2009), and reflect the limits of studies using only one proxy, particularly a financial indicator, to approximate problem gambling. Our tool is built in complete transparency. All variables that constituted the predictive model are easily extractable from online gambling providers' databases. Our model was designed to be used as an automatic screening instrument to trigger an intervention for controlled gambling. Moreover, our methods will allow the development of a more complete tool relevant for all online gambling types. Knowing the poor and late access to care of problem gamblers (Gainsbury *et al.,* 2014), our results open a way to address the need to enhance prevention strategies and particularly independent-from-gambling-providers ones. Playscan and other tools require a personal initiative to screen one's gambling behavior, and an agreement of the gambling provider. In France, regulation authorities stock routinely all account-based gambling data. Our tool and derived-following-our-methods future tools could be used by the French regulation authorities with no necessary agreement of gambling providers, and then allow to propose help to problem gamblers. It would allow authorities to resolutely engage in responsible gambling and assume their public health promotion mission. Our tool could instead provide information to problem gamblers who have not taken any

precaution before starting their gambling practice, which is the case for most gamblers. Independency from gambling providers of the screening method could keep away from any, even unintentional, conflict of interest or ambiguity in the screening purpose proposed by gambling providers. The recruitment was not only large but also systematic and devoid of incentives to avoid any recruitment bias. The problem gambling rate was congruent with the previously estimated prevalence of online problem gambling in France (Tovar *et al.,* 2013) (38% versus 31% of active online gamblers). The slightly higher rate could be due to lower socially desirable bias or to the inclusion of gamblers with confirmed accounts opened for at least one month.

Our study also presents several limitations. We could not document the refusal reasons or the reasons why gamblers did not open the study invitation email. Possible reasons are that the invitation email was detected as spam and thus automatically discarded or that it was considered a promotional email from Winamax. One significant limitation is the technical issue of documenting the gambling duration, even though we could document the number of gambling sessions. As previously explained, the Winamax application for wireless devices does not automatically disconnect after use. This issue occurred for only 57 gamblers and allowed for the interpretation of the entire sample. However, in consideration of the ease of use and to avoid errors, we chose not to include these variables in the predictive model. Although multi-tabling was included in the model and documented gamblers' tendency to roll on several games simultaneously, the model was built to be used by a single online gambling provider and only for poker gambling. Some gamblers could be involved in other gambling activities, and poker gambling may not be their most problematic mode of gambling as it has been shown for many subjects in a large study (Hing *et al.,* 2015). Some gamblers could be gambling on multiple gambling service providers' sites, where having several accounts is known to be more at risk of problem gambling (Gainsbury *et al.,* 2015). We also had no information regarding offline gambling practice. An additional possible limit is that our findings rely on the assumption that each account is used by only one person as it is legally supposed to in France. Further explorations should be conducted to reproduce our method, to develop a more comprehensive model if used by a regulation institution with access to data from several providers and several gambling modes. One possible use of our method could be the regular identification of at-risk gamblers by regulators, which could invite gambling providers to use less aggressive marketing interventions for this category of vulnerable gamblers, or even propose problem gamblers therapeutic interventions as normative feedback.

The specificity is quite low, approximately 50%, but can be explained by our choice of a conservative cutoff of the PGSI. The false positive rate decreases from 50% to 10% when considering only the not-at-risk population as false positives. We consider this low specificity a minimal drawback outweighed by the benefit of targeting at-risk risk populations for a preventive intervention. However the severity threshold, relevant to propose an intervention, meaning one needs or could actually benefit from help and support, is unclear and could be discussed. Taking into account the high level of harm for the problem gambler, his relatives and for society, we consider that our model is an acceptable risk reduction tool. The impact of a false warning in non-problem gamblers by the automatic screening process is however unknown and remains to be explored, for instance at the economic level for the operators in term of gambling participation.

Conclusion

We built and validated a model predictive of the PGSI status to be used as a tracking instrument for online poker problem gamblers, with good sensitivity of 80% and an acceptable rate of 10% of all gamblers who were not-at-risk gamblers but were misclassified as problem gamblers. This method should be reproduced and be used *in situ* to track problem online poker gamblers with no additional input from the gamblers. This automatic screening instrument allows for the implementation of targeted therapeutic interventions or preventive messages. Our method gives regulators and providers the keys to take responsibility for these vulnerable gamblers, who could now be identified. The study led to a partnership with French Authorities (Autorité de Régulation des Jeux En Ligne, ARJEL), who routinely collect all providers PABGD; our method is planned to serve as a basis to develop a more comprehensive instrument taking into account all providers PABGD and gambling modes.

Acknowledgements

Key points

- Our instrument can be used *in situ* to track problem online poker gamblers on Winamax website, with no additional input from the gamblers but the player account-based dataset.

- The development and validation of this instrument illustrates the feasibility of our method.
- Our method could be used to develop a more comprehensive instrument taking into account the several gambling service providers and the different gambling modes. Such an instrument would allow problem gamblers tracking to be routine by gambling service providers or regulators, as a prevention assessment.
- Our study led to a partnership with French authorities who routinely collect all providers player account-based gambling data; our method is planned to serve as a basis to develop a more comprehensive instrument taking into account all providers player account-based gambling data and gambling modes.

Declaration of interest statement

References

Assemblée Nationale (2010) LOI no. 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. NOR: BCFX0904879L. A. n. e. Sénat. 26, Paris, Assemblée Nationale.

Auer M., Griffiths M.D. (2013) Voluntary limit setting and player choice in most intense online gamblers: an empirical study of gambling behaviour. *Journal of Gambling Studies*, **29**, 647–660.

Barrault S., Varescon I. (2013) Cognitive distortions, anxiety, and depression among regular and pathological gambling online poker players. *Cyberpsychology, Behavior and Social Networking*, **16**(3), 183–188.

Braverman J., LaPlante D.A., Nelson S.E., Shaffer H.J. (2013) Using cross-game behavioral markers for early identification of high-risk internet gamblers. *Psychology of Addictive Behaviors*, **27**(3), 868–877.

Broda A., LaPlante D.A., Nelson S.E., LaBrie R.A., Bosworth L.B., Shaffer H.J. (2008) Virtual harm reduction efforts for Internet gambling: effects of deposit limits on actual Internet sports gambling behavior. *Harm Reduction Journal*, **5**, 27.

Canadian Consortium for Gambling Research. (2013) http://www.ccgr.ca/canadian-problem-gambling-index-cpgi

Costes J.M., Pousset M., Eroukmanoff V., Le Nezet O., Richard J.B., Guignard R., Beck F., Arwidson P. (2011) Les niveaux et pratiques des jeux de hasard et d'argent en 2010. *Tendances*, **77** (Septembre), 1–8.

Currie S.R., Casey D.M., Hodgins D.C. (2010) Improving the Psychometric Properties of the Problem Gambling Severity Index, Ottawa: Canadian Consortium for Gambling Research.

Delfabbro P., King D.L., Griffiths M. (2012) Behavioural profiling of problem gamblers: a summary and review. *International Gambling Studies*, **12**(3), 349–366.

Ferris J., Wynne H. (2001) The Canadian Problem Gambling Index: Final Report, Ottawa: Canadian Centre on Substance Abuse (CCSA).

Fink A., Parhami I., Rosenthal R.J., Campos M.D., Siani A., Fong T.W. (2012) How transparent is behavioral intervention research on pathological gambling and other gambling-related disorders? A systematic literature review. *Addiction*, **107**(11), 1915–1928.

Gainsbury S. (2011) Player account-based gambling: potentials for behaviour-based research methodologies. *International Gambling Studies*, **11**(2), 153–171.

Gainsbury S., Hing N., Suhonen N. (2014) Professional help-seeking for gambling problems: awareness, barriers and motivators for treatment. *Journal of Gambling Studies*, **30**(2), 503–519.

Gainsbury S.M., Russell A., Blaszczynski A., Hing N. (2015) Greater involvement and diversity of Internet gambling as a risk factor for problem gambling. *European Journal of Public Health*, **25**(4), 723–728.

Griffiths M., Whitty M. (2010) Online behavioural tracking in internet gambling research: ethical and methodological issues. *International Journal of Internet Research Ethics*, **3**(12), 104–117.

Griffiths M., Wardle H., Orford J., Sproston K., Erens B. (2009a) Sociodemographic correlates of internet gambling: findings from the 2007 British gambling prevalence survey. *CyberPsychology Behavior*, **12**(2), 199–202.

Griffiths M.D., Wood R.T., Parke J. (2009b) Social responsibility tools in online gambling: a survey of attitudes and behavior among Internet gamblers. *CyberPsychology Behavior*, **12**(4), 413–421.

Hing N., Russell A.M., Gainsbury S.M., Blaszczynski A. (2015) Characteristics and help-seeking behaviors of Internet gamblers based on most problematic mode of gambling. *Journal of Medical Internet Research*, **17**(1e13).

Hopley A.A., Nicki R.M. (2010) Predictive factors of excessive online poker playing. *Cyberpsychology, Behavior and Social Networking*, **13**(4), 379–385.

LaBrie R., Schaffer H.J. (2011) Identifying behavioral markers of disordered internet sport gambling. *Addiction Research and Theory*, **19**(1), 56–65.

LaPlante D.A., Kleschinsky J.H., LaBrie R.A., Nelson S.E., Shaffer H.J. (2009) Sitting at a virtual poker table: a prospective epidemiological study of actual Internet poker gambling behavior. *Computers in Human Behavior*, **25**(3), 711–717.

LaPlante D.A., Nelson S.E., Gray H.M. (2014) Breadth and depth involvement: understanding Internet gambling involvement and its relationship to gambling problems. *Psychology of Addictive Behaviors*, **28**(2), 396–403.

Miller N.V., Currie S.R., Hodgins D.C., Casey D. (2013) Validation of the problem gambling severity index using confirmatory factor analysis and Rasch modelling. *International Journal of Methods in Psychiatric Research*, **22**(3), 245–255.

National Council on Problem Gambling (2012) Internet Responsible Gambling Standards, Washington, DC: National Council on Problem Gambling.

Piazza P.V., Deroche-Gamonet V. (2013) A multistep general theory of transition to addiction. *Psychopharmacology (Berlin)*, **229**(3), 387–413.

Reilly C., Smith N. (2013) The Evolving Definition of Pathological Gambling in the DSM-5, Washington, DC: National Center for Responsible Gaming.

Tang Y.Y., Posner M.I., Rothbart M.K., Volkow N.D. (2015) Circuitry of self-control and its role in reducing addiction. *Trends in Cognitive Sciences*, **19**(8), 439–444.

Tovar M.L., Costes J.M., Eroukmanoff V. (2013) Les jeux d'argent et de hasard sur internet en France en 2012. *Tendances*, **85**, 1–6.

Wardle H., Moody A. et al. (2010) British Gambling Prevalence Survey, London: National Center for Social Research.

Williams R., Volberg R., Stevens R. (2012) The Population Prevalence of Problem Gambling: Methodological Influences, Standardized Rates, Jurisdictional Differences, and Worldwide Trends, Ontario: Ontario Problem Gambling Research Centre.