

Hospital de Clínicas de Porto Alegre
Universidade Federal do Rio Grande do Sul
Centro Colaborador em Álcool e Drogas HCPA/SENAD
Centro de Pesquisa em Álcool e Drogas

**ANÁLISES AVANÇADAS DE DADOS BRASILEIROS SOBRE DROGAS COM
INTELIGÊNCIA ARTIFICIAL E TRANSLAÇÃO PARA A CLÍNICA**

EQUIPE DE PESQUISA:

Investigador Principal: Prof. Dr. Flavio Pechansky

Coordenação do estudo: Dr. Felix Kessler e Dra. Lisia von Diemen

Porto Alegre, dezembro de 2017

1. Introdução

Ao longo da última década, o Centro de Pesquisa em Álcool e Drogas – CPAD/HCPA/UFRGS tem desenvolvido inúmeros projetos de ensino e pesquisa em parceria com órgãos nacionais, internacionais e, em especial, com a Secretaria Nacional de Políticas sobre Drogas – SENAD/MJ. Este histórico levou à implantação, a partir de 2012, do primeiro *Centro Colaborador com fins de Assistência, Ensino e Pesquisa em Álcool e outras Drogas* do HCPA/SENAD, estruturado de forma colaborativa entre as duas instituições na Unidade Álvaro Alvim do HCPA (convênio nº 24/2010). O principal objetivo deste Centro é desenvolver um modelo integrativo de cuidado e reinserção social de pacientes em nível ambulatorial e internação, totalmente incorporados a uma estrutura funcional de pesquisa, como é esperado de um centro colaborador de excelência. Além disso, o Centro Colaborador HCPA/SENAD tornou-se referência no país em estudos sobre álcool, drogas e trânsito através do pioneirismo em projetos de abrangência nacional e desenvolvimento de novas tecnologias nessa área. Visando maior integração com a sociedade, a partir do resultado destas produções, espera-se estimular o desenvolvimento de estudos sistemáticos para a disseminação de evidências sólidas sobre abuso de substâncias e comportamentos aditivos no país.

Dentre os diversos projetos de pesquisa já desenvolvidos destacam-se:

1. Projetos na área de assistência em álcool e drogas

- **Validação do Addiction Severity Index (ASI) versão 6 para o Brasil:** Este projeto foi realizado em parceria com a SENAD e com outros quatro centros de pesquisa de outras regiões do país. O ASI é uma escala de gravidade de dependência utilizada mundialmente, que foi desenvolvida em 1979 por Thomas McLellan e permite uma avaliação detalhada do uso de álcool e outras drogas. Este estudo produziu inúmeros dados como a própria versão brasileira do instrumento e seu manual, assim como artigos científicos abordando variados temas como comorbidades psiquiátricas, problemas sociais e familiares, entre outros, dos mais de 700 pacientes que foram avaliados. Ainda hoje, o banco de dados composto por essa amostra continua sendo explorado para melhor entender os fatores biopsicossociais envolvidos nos Transtornos relacionados ao uso de Substâncias Psicoativas, assim como para basear o desenvolvimento de novas estratégias preventivas e de tratamento.

- **Projeto Ações Integradas:** este projeto multicêntrico financiado pela SENAD foi desenvolvido em 5 capitais brasileiras e respectivas regiões metropolitanas: Porto Alegre – RS, Vitória – ES,

Salvador – BA, Rio de Janeiro – RJ e Brasília – DF, e abarcou cerca de 60 municípios. O objetivo foi fortalecer a rede de proteção social e desenvolver técnicas de prevenção e tratamento da dependência causada pelo uso de crack, álcool e outras drogas. Para tal, foram estabelecidos três eixos de ação: 1) mapeamento e georreferenciamento de instituições de atendimento a usuários de substâncias psicoativas; 2) cursos de capacitação para públicos específicos (saúde, assistência social, segurança pública e policiais rodoviários federais); 3) pesquisas com foco prioritário no consumo de crack, visando construir estratégias para tratamento e reinserção de usuários destas substâncias.

- **Projeto Preditores clínicos biológicos e psicossociais da recaída precoce em usuários de crack e álcool:** Financiado pelo CNPq e por recursos SENAD e FIPE/HCPA, este projeto permitiu o estabelecimento de um fluxo contínuo de coletas desde 2012. As coletas ocorrem na unidade de internação especializada em dependência química da Unidade Álvaro Alvim-HCPA/UFRGS, e todos os pacientes que ingressam no programa de tratamento são convidados a participar. O objetivo principal é avaliar fatores clínicos, biológicos e psicossociais e sua associação com recaída ou abandono de tratamento precoce em pacientes usuários de drogas após internação. Busca, ainda, verificar se há associação entre os níveis de BDNF, estresse oxidativo e NPY na internação e após desintoxicação e as taxas de recaída ou abandono de tratamento 30 dias após alta hospitalar; avaliar a associação de variáveis demográficas, comorbidades psiquiátricas e transtornos de personalidade com recaída precoce ou abandono precoce de tratamento; avaliar a associação de outras variáveis como impulsividade, trauma na infância, com a gravidade de dependência e recaída ou abandono de tratamento 30 dias após alta hospitalar e com variáveis neurobiológicas; avaliar a influência de polimorfismos genéticos em relação a gravidade da dependência e recaída ou abandono de tratamento 30 dias após alta hospitalar.

- **Ensaio Clínico Randomizado, Duplo-Cego, Controlado com Placebo, para Avaliar o Efeito da N-Acetilcisteína no Tratamento dos Transtornos por Uso de Álcool e Cocaína:** Este estudo financiado por verba SENAD encontra-se em fase de coleta de dados. Tem o objetivo de verificar se a N-acetilcisteína (NAC) aumenta a adesão ao tratamento quando usada como terapia adjuvante ao tratamento convencional dos transtornos por uso de álcool ou cocaína/crack. Baterias psicológicas e modificações nos níveis séricos de neurotrofinas, marcadores inflamatórios e marcadores de estresse oxidativo estão sendo coletados de forma

se avaliar se há associação com a resposta terapêutica com o uso de NAC como terapia adjuvante.

- **Projeto Cocaínas fumáveis na Argentina, Brasil, e Uruguai.** Este estudo financiado pela Organização dos Estados Americanos (OEA) concentra-se na investigação dos aspectos neuropsicológicos e neuroanatômicos em usuários de crack, a partir de uma amostra recrutada em 3 países da América do Sul, possibilitando uma visão ampliada da problemática. No Brasil, o Centro de Pesquisa em Álcool e Drogas - CPAD/HCPA/UFRGS é responsável pela coordenação. Com este projeto espera-se contribuir com um maior conhecimento sobre os mecanismos envolvidos na disfunção cognitiva e neurobiológica em indivíduos usuários de crack e avaliar possíveis níveis de gravidade da doença ao longo de um período da vida do usuário, os quais apresentam diferentes condições clínicas e respostas aos tratamentos hoje disponíveis. Com isso, futuramente embasar políticas públicas que oriente um tratamento mais direcionado ao perfil destes pacientes, podendo aumentar as taxas de adesão ao tratamento e diminuir a chance de recaída precoce.

2. Projetos na área de álcool, drogas e trânsito

- **Impacto do Uso de Bebidas Alcoólicas e Outras Drogas nas Rodovias Brasileiras:** maior estudo latino-americano já realizado sobre o tema, teve como objetivo estimar a prevalência de alcoolemia e de substâncias psicoativas entre motoristas de rodovias federais nas regiões metropolitanas das capitais brasileiras e do DF, realizar estudos de tecnologia avançada de aferição do custo econômico dos acidentes de trânsito, comorbidades psiquiátricas associadas ao ato de dirigir sob efeito de substâncias, além de geoprocessamento de acidentes relacionados a locais de venda e consumo de bebidas alcoólicas.

- **Tecnologias de Screening de SPAs no Trânsito:** este projeto, em sua fase inicial financiado pela SENAD, teve como objetivos principais: 1) avaliar tecnologias para detecção de substâncias psicoativas em condutores brasileiros; 2) descrever criticamente as tecnologias disponíveis e utilizadas pelas polícias em diversos países para detecção do uso de SPAs em condutores quanto à sua aplicabilidade, benefícios esperados e custos de implementação; 3) apresentar e discutir com profissionais e gestores públicos na área de trânsito as tecnologias de detecção de SPAs viáveis de serem implementadas no Brasil; 4) implementar, por meio de estudos pilotos com as polícias brasileiras, a utilização da(s) tecnologia(s) mais adequadas ao

contexto nacional. Para isso, durante a fase de coleta de dados, foram testados quatro diferentes equipamentos de obtenção de fluido oral (saliva) para aferição da presença de substâncias psicoativas em condutores, além amostras biológicas para a realização de testes confirmatórios em laboratório especializado.

- **Projeto Vida no Trânsito:** Coordenado pelo Ministério da Saúde, em uma articulação interministerial e parceria com a Organização Pan Americana da Saúde (OPAS), o Projeto Vida no Trânsito foi lançado em 2010 como parte da iniciativa internacional denominada *Road Safety in Ten Countries (RS-10)* sob a coordenação da *Johns Hopkins Bloomberg School of Public Health* e com verbas da *Bloomberg Philantropies*. Este projeto teve como objetivo realizar a vigilância e a prevenção de lesões e mortes no trânsito associadas principalmente ao excesso de velocidade e uso de álcool, em resposta aos desafios da Organização das Nações Unidas (ONU) para a Década de Ações pela Segurança no Trânsito 2011 – 2020. O Núcleo de Estudos e Pesquisa em Trânsito e Álcool (NEPTA) do CPAD/HCPA/UFRGS foi indicado pela SENAD para coordenar e executar a coleta de dados deste projeto em cinco cidades: Belo Horizonte; Campo Grande; Curitiba; Palmas e Teresina. As coletas foram realizadas a partir de dois objetivos: mensurar a velocidade de veículos em vias públicas (urbanas e federais) e obter informações de motoristas acerca de fatores de risco para colisões, incluindo o uso de álcool e outras drogas.

- **Projeto com o DETRAN-RS:** O DETRAN-RS coleta sistematicamente uma quantidade massiva de dados sobre trânsito no estado do Rio Grande do Sul, gerando bancos de dados completos sobre informações referentes a motoristas infratores. Desde 2012, o DETRAN-RS e o NEPTA/HCPA firmaram um termo de cooperação (acordo de cooperação N° 12/2012) para viabilizar o desenvolvimento de estudos sobre segurança no trânsito. Recentemente, o DETRAN-RS disponibilizou ao NEPTA uma base de dados contendo todos os casos de infrações por beber e dirigir da região metropolitana de Porto Alegre – totalizando aproximadamente 40.000 ocorrências. Apesar do recurso investido para a coleta e informatização desses dados, as informações sobre motoristas infratores são raramente avaliadas de forma conjunta e disponibilizadas para a sociedade.

Todos esses projetos geraram e continuam gerando dados de grande magnitude, com uma produção de bases de dados gigantes e com abrangência nacional em várias delas. Essas

bases foram analisadas até o momento através de métodos estatísticos tradicionais em seus projetos originais, mas possuem potencial para a utilização de ferramentas inovadoras.

Por que utilizar tecnologias inovadoras nesta área?

Nas últimas décadas houve um enorme avanço na área da informática com computadores com alta velocidade de processamento, memória de trabalho e capacidade de armazenamento. A disponibilidade dessa tecnologia permitiu um salto de conhecimento nas ciências exatas, em particular nas ciências da computação. A inteligência artificial e o aprendizado de máquina (“machine learning”) hoje são uma realidade no nosso dia a dia. Quando nos deparamos com sugestões de filmes ou séries no Netflix® ou com sugestão de amigos no Facebook®, as técnicas de aprendizado de máquina estão sendo aplicadas. Na Medicina, apesar de alguns rudimentos dessas técnicas estarem presente já na década de 60(1,2), a real utilização só começou de fato a partir dos anos 2000 (3–5) com o aprimoramento tecnológico. Na áreas da adições, o uso das técnicas de aprendizado de máquina é ainda incipiente, mas com resultados promissores(6,7): por exemplo, em um dos primeiros estudos na área, o uso dessa abordagem foi capaz de prever com 77% de acerto os indivíduos abstinentes de álcool em 12 semanas, em comparação com 42% em uma abordagem tradicional (8). A aplicação do aprendizado de máquina tem ocorrido também em áreas como álcool, drogas e trânsito, mas também de forma preliminar (9,10). Dessa forma, através do uso de simuladores, o uso de tais técnicas conseguiu identificar com 80% de acerto os motoristas que estavam sob efeito de álcool (10). Em resumo, as técnicas de aprendizado de máquina já trouxeram enormes avanços em áreas da medicina, como na radiologia (11) e na oncologia (12) e possuem um imenso potencial de avanço em áreas como psiquiatria, transtornos por uso de substâncias e uso de álcool e drogas no trânsito.

Qual a diferença entre a estatística tradicional e os métodos de aprendizado de máquina?

A estatística tradicional produz resultados que são baseados principalmente nas diferenças entre grupos de indivíduos, como por exemplo na avaliação de fatores associados à adesão ao tratamento de usuários de cocaína. Na abordagem tradicional, pode-se identificar que indivíduos mais jovens, desempregados, com doença psiquiátrica mais grave ficavam menos tempo em tratamento (14). Com o uso de técnicas de inteligência artificial é possível detectar padrões de características de indivíduos que ficam mais tempo em tratamento e desenvolver modelos preditivos para a adesão ao tratamento. Dessa forma, em vez de o resultado da análise ser relacionado a de não adesão, poderia ser desenvolvido um modelo

matemático preditivo de adesão ao tratamento. Esse modelo preditivo é desenvolvido “aprendendo” os padrões dos indivíduos que previamente aderiram ao tratamento, por isso o nome da técnica ser “aprendizado de máquina” ou inteligência artificial. Com modelos bem desenvolvidos, poderia ser possível calcular a chance de adesão individual de um determinado paciente. Na medida em que aumentamos os bancos de dados e fornecemos mais informações ao modelo, as predições irão se tornando mais precisas. Há uma percepção na comunidade científica de que estes métodos novos de análise de dados irão revolucionar a forma de fazer pesquisa, assim como já revolucionaram a indústria de marketing.

Dentro da perspectiva de ser um polo de inovação e de vanguarda na pesquisa em álcool e drogas, alguns pesquisadores do Centro Colaborador estão estudando e iniciando a aplicação das técnicas de inteligência artificial nos dados em coleta. Como já descrito, o Centro Colaborador produziu diversas bases de dados já analisadas de forma tradicional. Esses mesmos dados podem ser utilizados para análises de inteligência artificial com potencial de produzir dados que podem ter grande impacto em termos de saúde pública, em particular com atendimento de usuários de crack e de motoristas infratores cujas bases de dados são provavelmente as maiores do Brasil. Entretanto, as técnicas de inteligência artificial exigem um extenso trabalho de análise, necessitando de profissionais altamente qualificados e treinados para desenvolver e testar modelos preditivos para diferentes desfechos em saúde. Se por um lado há a exigência de alto investimento em pessoal qualificado, há uma economia de recursos em coleta de dados e uma otimização do uso dos recursos públicos já investidos em coletas de dados prévias.

Considerando este contexto, a realização deste projeto poderá contribuir, a médio e longo prazos, com a instrumentalização de profissionais para a construção de modelos inovadores de tratamento de usuários de crack, álcool ou outras drogas em um país de dimensões continentais como o Brasil. Espera-se, assim, real diminuição deste grave problema de saúde pública, em acordo com o que está previsto na Política Nacional sobre Drogas. Da mesma forma, os novos estudos poderão projetar a produção científica nacional de forma mais abrangente ao promover o intercâmbio com instituições regionais, nacionais e estrangeiras, além de organizações multinacionais similares. Isto permitirá aperfeiçoar uma rede colaborativa consistente com a Política Nacional e os compromissos hemisféricos e mundiais assumidos pelo país.

2. Objetivo

Avaliar, a partir de banco de dados já coletados pelo Centro Colaborador, preditores de adesão ao tratamento em usuários de substâncias assim como preditores de risco para beber e dirigir e conduzir com excesso de velocidade através de técnicas avançadas de análise estatística e computacional (Inteligência Artificial), incluindo medidas biológicas e variáveis sociodemográficas, psicossociais, comorbidades psiquiátricas, traumas na infância, gravidade do uso de substâncias, entre outras.

3. Objetivos Específicos

- Unificar bases de dados de usuários de crack e outras drogas de diferentes projetos propiciando análise conjunta dos indivíduos de diferentes procedências.

- Formar uma equipe com expertise para realizar análises de inteligência artificial na área de álcool e drogas, tornando o Centro Colaborador referência nacional na área.

4. Método

Este projeto propõe a unificação de bases de dados com foco na aplicação de técnicas estatísticas e computacionais na área de psiquiatria de adição. Utilizará dados secundários de bancos já coletados de amostras clínicas de usuários de substâncias provenientes de diferentes populações (CAPS AD, internações, ambulatório). Além disso, utilizará bancos de dados sobre o trânsito no Brasil coletados pelo Núcleo de Estudos e Pesquisa em Trânsito e Álcool do CPAD, e outras bases de dados com as quais o grupo vem trabalhando em parceria com outras instituições.

4.1. Bancos de Dados

Os Bancos de dados que serão utilizados para análise neste estudo abrangem resultados de diversos projetos realizados em parceria com o Centro de Pesquisa em Álcool e Drogas, Centro Colaborador do HCPA. Estes projetos incluem pacientes com transtornos por uso de substância, principalmente cocaínicos (inalada e fumada), álcool e cannabis, assim como, motoristas envolvidos ou não com infrações no trânsito. Estes bancos foram coletados ao longo dos últimos 10 anos em diversas investigações que tiveram parte dos resultados

publicados em algumas das principais revistas científicas da atualidade. Contudo, enfatizamos a premente necessidade de explorarmos de forma mais aprofundada e com técnicas mais atuais e avançadas, a fim de esclarecer questões em aberto e desvendarmos novos paradigmas para prevenção e tratamento.

Banco de dados na área de assistência em álcool e drogas

Os bancos de dados unificados somam um total de 3038 casos, dos quais 2261 possuem instrumentos coletados e 836 possuem amostras de sangue passíveis de análises biológicas. Os bancos de dados a serem analisados serão descritos de forma resumida abaixo:

- Banco Estudo de validação do Addiction Severity Index (ASI-6)

Neste estudo foram recrutados mais de 700 sujeitos, de ambos os sexos, com diagnóstico de dependência de cocaínicos - cocaína inalada e crack -, canabis e álcool. Transtornos psiquiátricos foram mensurados a partir da aplicação do Mini International Neuropsychiatric Interview (MINI) e a gravidade da dependência de drogas foi mensurada pelo Addiction Severity Index (Asi-6).

1. Banco Projeto Ações Integradas

Esta investigação avaliou mais de 750 sujeitos com diagnóstico de dependência de crack de ambos os sexos recrutados em 5 CAPS localizados em Porto Alegre, Rio de Janeiro, Salvador, Brasília, Vitória e São Paulo e 2 internações psiquiátricas de Porto Alegre, além disso, 216 controles saudáveis recrutados em um bairro periférico da região metropolitana de Porto Alegre. Transtornos psiquiátricos foram mensurados a partir da aplicação do Mini International Neuropsychiatric Interview (MINI) e a gravidade da dependência de drogas foi mensurada pelo Addiction Severity Index (Asi-6), entre outros instrumentos, como o questionário de trauma na infância, de impulsividade e de estilos parentais. Parte desta amostra de casos e controles tiveram amostras de sangue coletadas, e a dosagem de biomarcadores foi realizada - BDNF, interleucinas, estresse oxidativo, NPY, e doenças infectocontagiosas - uma parcela das amostras biológicas permanecem armazenadas a - 80 °C.

- Banco Estudos com adolescentes

Trata-se de um subprojeto do Ações integradas com enfoque em adolescentes. 90 adolescentes de 12 a 17 anos internados por dependência de crack, todos tiveram o Addiction Severity Index (T-ASI) coletado e amostra de sangue coletadas para avaliação da IL-6, IL-10, TBARS e BDNF.

- **Banco Projeto Preditores clínicos biológicos e psicossociais da recaída precoce em usuários de crack e álcool**

Este banco integra coletas realizadas desde 2012. Até dezembro de 2017 cerca de 1038 sujeitos do sexo masculino foram recrutados, o que resultou em um banco de dados formado por 541 usuários de crack, 81 usuários de cocaína, 294 usuários de álcool e 122 usuários de outras drogas principalmente maconha, inalantes, sedativos e club drugs. Todos os sujeitos possuem exames clínicos e a maioria teve as sorologias realizadas. Em decorrência de altas precoces nem todos os sujeitos completaram a bateria de coletas, atualmente o banco de dados concentra: 671 protocolos do Addiction severity Index – ASI 6; 591 Entrevistas Clínicas Estruturadas para os transtornos do eixo I do DSM-IV (SCID I); 91 Entrevistas Clínicas Estruturadas para os transtornos do eixo II (SCIDS II); 480 Questionários de trauma na infância; 202 questionários de resiliência; 554 Barrats que avalia impulsividade; 491 amostras de sangue coletadas na baixa; 343 amostras de sangue coletadas na alta.

- **Banco Ensaio Clínico Randomizado, Duplo-Cego, Controlado com Placebo, para Avaliar o Efeito da N-Acetilcisteína no Tratamento dos Transtornos por Uso de Álcool e Cocaína**

Este estudo o qual investiga a N-Acetilcisteína no tratamento da dependência de álcool e crack possui 27 pacientes coletados até o momento com avaliação de instrumentos psicológicos e amostras biológicas.

- **Banco Cocaínas fumáveis na Argentina, Brasil e Uruguai**

Iniciado em janeiro de 2017 com o objetivo de investigar alterações da função cerebral em usuários de crack e cocaína. O projeto prevê coleta de 50 sujeitos das quais 25 já foram realizadas. Estes dados incluem baterias de avaliação Psiquiátrica e Neuropsicológica além de exames de ressonância magnética funcional.

Bancos de dados sobre trânsito, álcool e drogas

- **Banco Projeto Vida no Trânsito**

Neste projeto foram gerados dois bancos de dados com uma quantidade significativa de casos, conforme descrito a seguir:

- I. Banco de dados de velocidade: contém mais de 460.000 observações de velocidade de veículos. Além dos dados de velocidade, esse banco contém

informações sobre o tipo de veículo, tipo de via, condições da via, presença de fiscalização e de sinalização e limites de velocidade.

- II. Banco de dados de fatores de risco: contém informações sobre cerca de 9.000 motoristas, que foram entrevistados em locais públicos nas cidades de interesse. Durante a entrevista, os motoristas forneceram informações sobre os comportamentos de risco no trânsito sabidamente associados com colisões, tais como excesso de velocidade, uso de álcool e drogas, uso de telefone celular e uso de sinto de segurança.

- Banco de dados Bases do DETRAN-RS

A base de dados disponibilizada ao NEPTA pelo DETRAN-RS totaliza aproximadamente 40.000 ocorrências de infrações por beber e dirigir (região metropolitana de Porto Alegre). Nesse sentido, a análise de bancos de dados coletados pelo DETRAN-RS permitirá a análise da totalidade dos casos de interesse, como por exemplo, motoristas autuados pelo uso de álcool, viabilizando a otimização dos recursos federais a fim de gerar informações sólidas e consistentes sobre o perfil do motorista infrator, até hoje desconhecido no Brasil.

- Banco de dados Tecnologias de Screening de SPAs no Trânsito

Este projeto possibilitou a coleta de dados de 178 motoristas abordados em barreiras de fiscalização na cidade de Porto Alegre, resultando em um banco de dados contendo informações sociodemográficas, perfil do motorista, histórico do uso de drogas, opiniões sobre métodos de fiscalização de trânsito, entre outros. Além disso, esse estudo permitiu a coleta de medidas de alcoolemia no momento da abordagem, bem como a coleta de amostras biológicas (fluido oral/saliva) para a avaliação de tecnologias de screening de drogas e avaliação toxicológica por métodos confirmatórios (análise cromatográfica).

5. Análise dos dados

Unificação e preparação dos bancos de dados: uma etapa importante para a realização das análises de inteligência artificial é a unificação do banco de dados e a preparação dos bancos para a realização das análises. Em um primeiro momento, é possível unificar bancos de dados em que tenham sido coletados dados através do mesmo instrumento ou de

instrumentos semelhantes. Por exemplo, a variável tentativa de suicídio pode vir de uma amostra que tenha ASI, ou de outra que tenha SCID ou ainda de outra que tenha sido coletado o MINI. Todas elas podem fornecer essa informação, mas para juntar em um mesmo banco de dados deve ser analisado se em todas tem o mesmo significado. E isso para cada uma das variáveis relevantes. Essa é uma parte extremamente trabalhosa, mas que permite a obtenção de resultados mais fidedignos, pois possibilita que se construam bancos de dados muito maiores do que cada projeto individualmente.

Análise exploratória univariada e bivariada: A aderência das variáveis quantitativas à distribuição Normal será investigada por meio de histograma e teste Shapiro-Wilk. As que evidenciarem normalidade ou não desviarem fortemente da simetria serão apresentadas por média e desvio padrão, correlacionadas por coeficiente de Pearson e comparadas entre grupos por teste T, ou entre mais de dois grupos por ANOVA. As que não evidenciarem normalidade ou apresentarem forte assimetria serão apresentadas por mediana e intervalo interquartil, correlacionadas por coeficiente de Spearman e comparadas entre grupos pelo teste não paramétrico Mann-Whitney, ou entre mais de dois grupos por Kruskal-Wallis. As variáveis categóricas serão apresentadas por frequência absoluta e relativa, associações serão investigadas via teste Qui-Quadrado e Razão de Prevalência. O nível de significância adotado será de 5%.

Modelos lineares generalizados (GLM): Para a modelagem dos dias de permanência na internação e do tipo de alta hospitalar, alguns GLM serão ajustados e comparados. Para dias de permanência, os modelos serão: regressão Poisson com função de ligação logarítmica e regressão Gama com funções de ligação logarítmica e recíproca. E para o tipo de alta hospitalar, os modelos serão: regressão Logística com função de ligação *logit* e a regressão Binomial com função de ligação logarítmica. Para a seleção das variáveis independentes será utilizado o método *stepwise*, com critérios de exclusão e inclusão em $p\text{-valor} > 0,1$ e $p\text{-valor} < 0,05$, em um grupo de variáveis pré avaliadas que não apresentem correlação espúria com as variáveis dependentes mencionadas. Para comparabilidade dos modelos, serão utilizadas as seguintes estatísticas de qualidade de ajuste: AIC, BIC, pseudo- R^2 e curva ROC.

Aprendizado de máquina ou inteligência artificial (*Machine-Learning*): Também para modelagem dos dias de permanência na internação e do tipo de alta hospitalar, serão utilizados algoritmos de aprendizagem supervisionada. Para dias de permanência serão aplicados: Árvores de decisão (*Decision Trees*; DT) , *Random Forest* (RF), *Ensemble Methods* (EM) e Redes Neurais Artificiais (RNA). E para o tipo de alta hospitalar serão aplicados: *Support Vector Machines* (SVM) e *Nearest Neighbour* (kNN). O banco de dados será dividido

aleatoriamente em dois, a primeira parte conterá 80% dos dados e será utilizada para a etapa de aprendizagem dos algoritmos de ML, a segunda conterá 20% e será utilizada para a etapa de teste dos modelos preditivos. Para eleição das variáveis de maior importância para predição será utilizado o método *Recursive Feature Elimination*, e para avaliação da reprodutibilidade e consistência dos modelos, o banco destinado ao processo de aprendizagem será particionado por *3-fold cross-validation*.

Uma alternativa de análise que também pode ser empregada é a modelagem da fase do tratamento em que o paciente estava no momento da alta hospitalar, isto é, a utilização da variável de grupo de desintoxicação (GD) ou reabilitação (GR). O armazenamento e análises dos dados serão realizados pelos softwares IBM SPSS versão 18 e R-project versão 3.4.1 com interface RStudio.

6. Metas

Meta 1 - Capacitar pesquisadores para análises de inteligência artificial na área de álcool e drogas.

- **Produto Meta 1** - Esta meta terá como produto final a redação de 1 (um) Manual para Pesquisadores sobre o uso de inteligência artificial (*Machine Learning*) em pesquisas sobre álcool e outras drogas (descrição completa no item 8. Produtos).

Banco de dados de amostras clínicas de usuários de substâncias

Meta 2 - Aplicar modelos preditivos de *Machine Learning* para adesão ao tratamento hospitalar entre usuários de crack e outras drogas possibilitando avaliações mais complexas e individualizadas.

Meta 3 - Testar os conceitos de estadiamento e de medicina personalizada em transtorno por uso de substâncias através de modelos que utilizem alterações de biomarcadores e características individuais ao longo da progressão da doença de forma translacional visando o desenvolvimento de novas estratégias preventivas e terapêuticas.

- **Produto Metas 2 e 3** – Estas metas terão como produto a redação de 1 (um) relatório parcial contendo a metodologia da pesquisa e resultados iniciais (descrição completa no item 8. Produtos).

Banco de dados do trânsito, álcool e drogas

Meta 4 - Aplicar algoritmos de *Machine Learning* para avaliar a predição de risco para beber e dirigir através de variáveis relacionadas ao padrão de direção, uso de substâncias psicoativas, características sociodemográficas e outras infrações de trânsito;

Meta 5 - Desenvolver um modelo preditivo através de técnicas de análise de *Machine Learning* para a análise de velocidade de veículos automotores em rodovias públicas brasileiras;

- **Produto Metas 4 e 5** – Estas metas terão como produto a redação de 1 (um) relatório parcial contendo a metodologia da pesquisa e resultados iniciais (descrição completa no item 8. Produtos).

Meta 6 – Interpretar e realizar a translação dos achados das análises realizadas para aplicação clínica baseada em tratamento personalizado e para tomadas de decisão acerca de políticas públicas.

- **Produto Meta 6** – Esta meta terá como produto a redação do relatório final contendo os principais resultados das análises e a interpretação e translação dos mesmos (descrição completa no item 8. Produtos).

7. Etapas

Para atingir as metas propostas pelo projeto, estão previstas 8 etapas:

Etapas 1: Seleção da equipe, escrita e submissão do projeto de pesquisa: nesta etapa serão selecionados os pesquisadores júnior e sênior e o projeto de pesquisa será escrito e submetido à Plataforma Brasil e ao comitê de ética em pesquisa do HCPA.

Etapas 2: Treinamento e supervisão da equipe em inteligência artificial: o consultor em inteligência artificial irá organizar um curso teórico para os pesquisadores do centro

colaborador e fará supervisão sistemática da unificação e organização dos bancos de dados, bem como da realização das análises dos dados.

Etapa 3: Unificação e organização dos bancos de dados: os coordenadores do projeto em conjunto com os pesquisadores selecionados irão planejar quais variáveis de quais bancos de dados podem ser unificados, definir quantos bancos de dados agregados serão possíveis de se obter e realizar a unificação. Cada pesquisador terá sob sua responsabilidade a construção de um ou mais bancos de dados unificados.

Etapa 4: Revisão de literatura: Será realizado uma revisão da literatura sobre modelos de estadiamento e de medicina personalizada em usuários de álcool e outras drogas, assim como os principais métodos de análise de grandes dados já publicados sobre o tema na literatura.

Etapa 5: Planejamento das análises: após a construção dos bancos de dados serão planejados os modelos de inteligência artificial que serão testados. Serão definidos para quais desfechos serão desenvolvidos modelos preditivos através de inteligência artificial e quais variáveis predictoras serão testadas no modelo.

Etapa 6: Execução das análises: a obtenção de um modelo preditivo através de inteligência artificial implica em tempo de análise maior do que nos métodos tradicionais, pois existem diversas técnicas diferentes e várias delas são testadas para se achar o melhor modelo preditivo.

Etapa 7: Translação dos resultados: após a obtenção de modelos preditivos através de inteligência artificial, há uma necessidade de transpor esses modelos para possíveis implicações na atenção dos usuários de substâncias ou na elaboração de políticas públicas, como no caso do trânsito. Nesta etapa será discutido e planejado como realizar o manual para pesquisadores para o uso de inteligência artificial na área de álcool e drogas e o relatório com os resultados para a SENAD com as possíveis aplicações práticas dos resultados obtidos.

Etapa 8: Elaboração do manual e relatório para a SENAD: a escrita do manual para pesquisadores e o relatório para a SENAD será dividida entre os coordenadores e pesquisadores do projeto.

8. Produtos

Espera-se que os resultados obtidos a partir da sistematização de dados do projeto possam colaborar com o entendimento de tendências e padrões de comportamento relacionados ao consumo de substâncias no contexto brasileiro. Para tal, foram definidos como produtos, além da capacitação da equipe, um relatório de pesquisa parcial, um relatório de pesquisa final, um

manual, que posteriormente poderão contribuir para o desenho de políticas públicas e estratégias de tratamento baseadas em evidências científicas:

- **Produto 1 – Relatório de Pesquisa Parcial:** Serão entregues 02 (dois) relatórios parciais: relatório parcial 1 e relatório parcial 2, respectivamente. Os relatórios apresentarão as diferentes análises realizadas e os principais resultados obtidos, que poderão contribuir com a validação de indicadores no campo das adições e o estabelecimento de abordagens adequadas a características específicas dos usuários. Visando gerar impacto no meio científico, a divulgação deste estudo posteriormente será compartilhada em periódicos nacionais e internacionais no formato de artigos específicos.

- **Produto 2 – Relatório de Pesquisa Final:** Será entregue 01 (um) relatório final. Os relatórios apresentarão as diferentes análises realizadas e os principais resultados obtidos, que poderão contribuir com a validação de indicadores no campo das adições e o estabelecimento de abordagens adequadas a características específicas dos usuários. Visando gerar impacto no meio científico, a divulgação deste estudo posteriormente será compartilhada em periódicos nacionais e internacionais no formato de artigos específicos.

- **Produto 3 – Manual para Pesquisadores sobre o uso de inteligência artificial (*Machine Learning*) em pesquisas sobre álcool e outras drogas:** tendo em vista que as análises com o uso de inteligência artificial têm o potencial de estabelecer preditores e contribuir com o aumento da acurácia diagnóstica, a experiência do uso desta técnica será transformada em 01 (um) manual para profissionais e pesquisadores. Este manual terá cunho didático e será produzido em uma linguagem técnica acessível, descrevendo passo a passo as etapas prévias necessárias para a realização de análises de grandes bases de dados (*big data*): unificação de bancos, definição de variáveis, testagem de modelos e possíveis análises, compreensão dos fatores subjacentes aos desfechos positivos em tratamentos. Espera-se que este manual sirva como ferramenta estratégica em pesquisas na área de drogas.

- **Produto 4 – Profissionais Capacitados na técnica de *Machine Learning*:** capacitação básica de 20 profissionais do Centro Colaborador em Álcool e Drogas do HCPA/SENAD para que adquiram maior expertise na utilização desta técnica, visando o desenvolvimento de pesquisas de plano analítico de maior complexidade. Estes profissionais poderão contribuir com o desenvolvimento de outros estudos, a disseminação de conhecimento e com a busca de informações diferenciadas e úteis para a prática clínica. Além destes, serão capacitados 05 (cinco) pesquisadores em nível avançado, para que estes tenham o domínio pleno da aplicação dos métodos de *Machine Learning* dentro da temática das adições.

Tabela 1. Quantitativo, data prevista de entrega e especificações dos produtos

Produtos	Unidade	Quant Prevista	Data prevista	Especificações do produto
Profissionais capacitados	Arquivo eletrônico	20 (básico) e 5 (avançado)	30/12/2018	Lista nominal dos pesquisadores capacitados, certificação aos participantes, relatório de execução do curso de capacitação
Relatório Parcial	Arquivo Eletrônico	2	31/07/2019	Relatório contendo a metodologia da pesquisa e resultados parciais das análises (relatório Parcial 1 e relatório Parcial 2)
Planejamento das análises	Arquivo Eletrônico	1	31/12/2019	Manual descrevendo a experiência com a metodologia empregada e possíveis aplicações práticas a ser disponibilizado em formato eletrônico para pesquisadores e profissionais da área
Manual para Execução das análises pesquisadores				
Revisão da literatura				
Relatório final das análises	Relatório Parcial 1 Arquivo Eletrônico	1	31/12/2020	Relatório contendo os principais resultados das análises e a interpretação e translação dos mesmos
9. Cronograma de Execução – Metas, Etapas, Produtos e Resultados				
Unificação e organização dos bancos de dados	Relatório Parcial 2	Relatório contendo a metodologia da pesquisa e resultados parciais	01/07/2018 a 31/07/2019	
Planejamento das análises				
Execução das análises				
Revisão da literatura	Relatório Parcial 2	Relatório contendo a metodologia da pesquisa e resultados parciais	01/07/2018 a 31/07/2019	
Planejamento das análises				
10. Considerações Éticas				
Meta	Etapas	Produtos	Resultado	Execução
Translação dos resultados	Seleção da equipe, escrita e submissão do projeto de pesquisa	Relatório contendo principais resultados	Lista nominal dos pesquisadores capacitados, certificação aos participantes, relatório de execução do curso de capacitação	29/12/2017 a 30/12/2018
Unificação dos Resultados e desenvolvimento do relatório final pesquisadores	Treinamento e supervisão da equipe em inteligência artificial	Profissionais capacitados e translação dos mesmos	Manual descrevendo métodos de <i>Machine Learning</i> e sua aplicação nas adições	
	Escrita do manual	Manual para pesquisadores	Manual descrevendo métodos de <i>Machine Learning</i> e sua aplicação nas adições	01/01/2019 a 31/12/2019

Testar os conceitos de estadiamento e medicina personalizada

Aplicar *Machine Learning* para avaliar a predição de risco para beber e dirigir

Desenvolver um modelo preditivo para a análise de velocidade de veículos

Interpretação e translação dos achados para aplicação clínica e em políticas públicas

Os dados para os bancos analisados foram retirados de projetos do (CPAD-HCPA/SENAD), já aprovados pelo Comitê de Ética e Pesquisa do Hospital de Clínicas de Porto Alegre. O presente projeto proposto não exige coleta de novos dados e todos os sujeitos de pesquisa assinaram o Termo de Consentimento Livre e Esclarecido referente aos respectivos projetos. Mesmo assim, esta proposta será submetida ao CEP-HCPA no formato de um novo projeto de pesquisa.

11. Equipe do Projeto

Coordenadores do Projeto: o projeto terá a coordenação dos professores Flavio Pechansky, Lisia von Diemen e Felix Kessler. Este último estará mais envolvido no dia a dia do andamento do projeto, requerendo uma carga horária semanal fixa para o projeto. Por esta razão, será solicitada uma bolsa de pesquisa apenas para ele.

1. Flavio Pechansky

Professor Titular do Departamento de Psiquiatria da UFRGS, Diretor do Centro de Pesquisa em Álcool e Drogas da UFRGS, Chefe do Serviço de Psiquiatria de Adição do Hospital de Clínicas de Porto Alegre e Coordenador do Centro Colaborador em Álcool e Drogas HCPA/SENAD.

Atribuições: Responsabilizar-se-á pela coordenação geral e implementação do projeto de pesquisa; coordenação geral da seleção da equipe; definição de diretrizes técnicas e executivas; revisar e responsabilizar-se pelos relatórios finais. Supervisionará a seleção de profissionais para o projeto.

2. Lisia von Diemen

Professora adjunta do Departamento de Psiquiatria da UFRGS, Vice-Diretora do Centro de Pesquisa em Álcool e Drogas da UFRGS, Chefe da Unidade de Ensino e Pesquisa do Serviço de Psiquiatria de Adição do Hospital de Clínicas de Porto Alegre.

Atribuições: Coordenar a seleção da equipe de pesquisa; Responsável pela elaboração final do projeto de pesquisa; Coordenação da elaboração do protocolo de pesquisa, incluindo as medidas utilizadas para avaliar as variáveis em estudo e os desfechos clínicos; Coordenar o treinamento da equipe de pesquisa para análise dos dados; Coordenar a confecção do banco de dados; Realizar reuniões semanais com a equipe do projeto para resolver problemas no andamento do projeto, organizar as etapas a seguir e supervisão do andamento do projeto;

Coordenar a elaboração dos relatórios técnicos para envio para a SENAD; Supervisionar sistematicamente a equipe de pesquisa durante a análise dos dados.

3. Félix Henrique Paim Kessler

Atualmente é Professor Adjunto do Departamento de Psiquiatria e Medicina Legal da UFRGS e Coordenador do Núcleo de Pesquisa Clínico-Biológica e Núcleo de Pesquisa em Neuroimagem e do Ambulatório de Psiquiatria de Adição do HCPA.

Atribuições: Coordenar a seleção da equipe de pesquisa; Responsável pela elaboração final do projeto de pesquisa; Coordenação da elaboração do protocolo de pesquisa, incluindo as medidas utilizadas para avaliar as variáveis em estudo e os desfechos clínicos; Coordenar o treinamento da equipe de pesquisa para análise dos dados; Coordenar a confecção do banco de dados; Realizar reuniões semanais com a equipe do projeto para resolver problemas no andamento do projeto, organizar as etapas a seguir e supervisão do andamento do projeto; Coordenar o planejamento da execução do projeto; Elaborar os relatórios técnicos para envio para a SENAD; Supervisionar sistematicamente a equipe de pesquisa durante a análise dos dados.

Equipe para o projeto:

Para a execução deste projeto serão necessários profissionais especializados na área de adição capazes de coordenar, revisar e analisar os bancos gerados através inteligência artificial.

As bolsas de pesquisa previstas visam unicamente viabilizar a execução do projeto, não tendo caráter de prestação de serviço permanente. Por tratar-se de um projeto com metas específicas, a ser executado por tempo determinado, o período de oferta das bolsas não excederá o máximo de 3 anos para o desenvolvimento das análises previstas.

Bolsa para coordenador técnico: o coordenador técnico será o Professor Felix Kessler, que será responsável pela seleção da equipe, organização do treinamento e capacitação da mesma, coordenação do planejamento das análises e supervisão da elaboração dos artigos e relatórios do projeto.

Bolsa de Pesquisa para Pesquisador sênior (3): A estrutura mínima da equipe de pesquisa prevê a participação de 03 profissionais com nível de doutorado. Estes profissionais serão treinados para aplicação de inteligência artificial nos bancos de dados de álcool e drogas. Inicialmente irão realizar a unificação e preparação dos bancos de dados para que essas

análises possam ser realizadas. Eles serão responsáveis por propor e realizar os modelos de análise, inicialmente com a ajuda do consultor e depois de forma independente. Terão responsabilidade final na produção de relatórios, materiais de divulgação e artigos científicos.

Bolsa de Pesquisa para Pesquisador Junior(1): Previsão de bolsa para 01 profissional com formação de nível superior, com experiência comprovada em projetos de pesquisa em áreas afins. Este profissional será responsável pela conferência, organização dos instrumentos de pesquisa e banco de dados. Participarão da elaboração de relatórios de pesquisa, artigos e *abstracts* para eventos científicos.

Bolsa para consultor em inteligência artificial(1): Este projeto prevê a realização de treinamento e consultorias na área de Inteligência Artificial para a equipe do Centro Colaborador, de forma a capacitar o grupo de pesquisa e garantir a produção sistemática de conhecimento. Para tanto, o Professor Ives Passos, que já trabalha com uso de inteligência artificial no transtorno bipolar, irá realizar o treinamento da nossa equipe e consultorias sistemáticas na implementação das análises de inteligência artificial por um período de 2 anos.

12. Materiais para o Desenvolvimento de Estudos e Aquisição de Equipamentos de Uso Permanente

Com vistas à estruturação do Centro e o melhor desenvolvimento dos estudos e análise de dados produzidos na Unidade de Adição, será necessária a aquisição de alguns equipamentos. É necessária a aquisição de equipamentos que viabilizem a preparação, processamento e armazenamento dos materiais biológicos coletados na Unidade Álvaro Alvim. São eles:

Ultra freezer:

A qualidade dos resultados dos exames laboratoriais está intimamente relacionada à fase pré-analítica, que se inicia desde a correta preparação do paciente, coleta, identificação da amostra, centrifugação do material, além do armazenamento da amostra no laboratório. Para tanto, faz-se imprescindível a aquisição de um Ultra Freezer (ou Ultra Congelador), a fim de manter continuamente um exato controle de temperatura e umidade das amostras coletadas. Amostras criogênicas e biológicas em geral, exigem determinados cuidados e o armazenamento sob baixas temperaturas, e isto é um requisito primordial para que posteriormente as análises sejam realizadas com o mínimo de imprecisão. O Ultra Freezer é empregado na conservação e armazenamento de materiais para laboratório e preparação de

amostras que serão submetidas a testes de qualidade, já que promove temperaturas baixas, variando de -50 a -86°C.

Até agora, o material coletado é transportado até os freezers existentes na sede do HCPA. Entretanto, a necessidade do transporte até o hospital frequentemente causa deterioração ou até mesmo perda do material biológico. As coletas realizadas na Unidade Álvaro Alvim atualmente não contam com esse dispositivo, o que reduz ou até mesmo inviabiliza determinadas pesquisas.

Centrífuga de laboratório:

As centrífugas de laboratório são empregadas na separação de amostras. Poucos experimentos são realizados em laboratórios sem ao menos uma etapa de centrifugação. São utilizadas para preparar produtos de análise celular, genético e de proteínas. O material é colocado em tubos de ensaio e alocado neste equipamento. Com a rotação, a parte sólida se separa da parte líquida. São parte fundamental no processo de aquisição de material biológico, visto que, após a coleta, é necessária a preparação do material para posterior armazenamento ou análise. Diante disso, faz-se necessária a aquisição deste equipamento para viabilizar a etapa de processamento das amostras.

13. Orçamento

Além das despesas com bolsas e material permanente, fica estabelecida a necessidade de repasse de um percentual estimado de até 05% (cinco por cento) sobre os valores das bolsas, visando à cobertura de custos administrativos de pessoal, insumos e serviços inerentes a execução das atividades deste Plano de Trabalho, estimado ao valor máximo de R\$ 35.380,00.

	Item	Quant.	Tempo	Unitário (R\$) Valor Bruto	Total
Bolsas*	Coordenador técnico	1	34	\$4.100,00	\$139.400,00
	Pesquisador sênior	3	34	\$4.100,00	\$418.200,00
	Pesquisador Junior	1	34	\$3.000,00	\$102.000,00
	Despesas Administrativa	1	1	\$50.453,00	\$50.453,00 ¹
	Consultoria em inteligência artificial	1	20	\$2.066,00	\$41.320,00
Equipamentos**					
	Centrífuga	1	1	\$57.000,00	\$57.000,00
	Freezer -80°C	1	1	\$75.000,00	\$75.000,00

Total					\$883.373,00
-------	--	--	--	--	--------------

*As bolsas de pesquisa estão previstas para iniciar em Março de 2018

**Os equipamentos deverão ser empenhados em Dezembro de 2017 para pagamento em Janeiro de 2018

¹ O ressarcimento das despesas administrativas terão o limite de até R\$ 35.380,00.

14. Cronograma de Desembolso

Desembolsos	Valor	Grupo de Despesa - GND	Detalhamento do GND
2017	R\$ 132.000,00	GND 449052	Material Permanente
2018	R\$ 214.660,00	GND 339048	Bolsas de Pesquisa/Extensão
2018	R\$ 14.838,00	GND 339039	Despesas Administrativas
2019	R\$ 253.460,00	GND 339048	Bolsas de Pesquisa/Extensão
2019	R\$ 15.410,00	GND 339039	Despesas Administrativas
2020	R\$ 232.800,00	GND 339048	Bolsas de Pesquisa/Extensão
2020	R\$ 20.205,00	GND 339039	Despesas Administrativas
Total Geral			R\$ 883.373,00

15. Contrapartida do proponente

Cabe salientar que verbas para a realização de outras atividades do projeto que não constam no orçamento acima apresentado, tais como apresentação de resultados em congressos, materiais de escritório, manutenção de espaço físico, entre outros, serão obtidos pelo proponente por outras fontes de financiamento e serão apresentados na prestação de contas do projeto, se houverem.



Flavio Pechansky
Coordenador do Projeto

16. Referências

1. Reed J, Toombs R, Barricelli NA. Simulation of biological evolution and machine learning. I. Selection of self-reproducing numeric patterns by data processing machines,

- effects of hereditary control, mutation type and crossing. *J Theor Biol.* 1967 Dec;17(3):319–42.
2. Findler N V. Machine learning from noisy information. *Nature.* 1964 Oct 3;204:103.
 3. Pearce CB, Gunn SR, Ahmed A, Johnson CD. Machine learning can improve prediction of severity in acute pancreatitis using admission values of APACHE II score and C-reactive protein. *Pancreatology.* 2006 Apr;6(1–2):123–31.
 4. Burgansky-Eliash Z, Wollstein G, Chu T, Ramsey JD, Glymour C, Noecker RJ, et al. Optical coherence tomography machine learning classifiers for glaucoma detection: a preliminary study. *Invest Ophthalmol Vis Sci.* 2005 Nov 1;46(11):4147–52.
 5. Nattkemper TW, Arnrich B, Lichte O, Timm W, Degenhard A, Pointon L, et al. Evaluation of radiological features for breast tumour classification in clinical screening with machine learning methods. *Artif Intell Med.* 2005 Jun;34(2):129–39.
 6. Acion L, Kelmansky D, van der Laan M, Sahker E, Jones D, Arndt S. Use of a machine learning framework to predict substance use disorder treatment success. Niaura R, editor. *PLoS One.* 2017 Apr 10;12(4):e0175383.
 7. Mete M, Sakoglu U, Spence JS, Devous MD, Harris TS, Adinoff B. Successful classification of cocaine dependence using brain imaging: a generalizable machine learning approach. *BMC Bioinformatics.* 2016 Oct 6;17(Suppl 13):357.
 8. Connor JP, Symons M, Feeney GFX, Young RM, Wiles J. The Application of Machine Learning Techniques as an Adjunct to Clinical Decision Making in Alcohol Dependence Treatment. *Subst Use Misuse.* 2007 Jan 3;42(14):2193–206.
 9. Chen H, Chen L. Support Vector Machine Classification of Drunk Driving Behaviour. *Int J Environ Res Public Health.* 2017 Jan 23;14(1):108.
 10. Li Z, Jin X, Zhao X. Drunk driving detection based on classification of multivariate time series. *J Safety Res.* 2015 Sep;54:61–4.
 11. Lakhani P, Prater AB, Hutson RK, Andriole KP, Dreyer KJ, Morey J, et al. Machine Learning in Radiology: Applications Beyond Image Interpretation. *J Am Coll Radiol.* 2017 Nov 16;
 12. Gupta S, Tran T, Luo W, Phung D, Kennedy RL, Broad A, et al. Machine-learning prediction of cancer survival: a retrospective study using electronic administrative records and a cancer registry. *BMJ Open.* 2014 Mar 17;4(3):e004007.
 13. Moeller FG, Dougherty DM, Barratt ES, Schmitz JM, Swann AC, Grabowski J. The impact of impulsivity on cocaine use and retention in treatment. *J Subst Abuse Treat.* 2001;21(4):193–8.

14. Siqueland L, Crits-Christoph P, Gallop R, Barber JP, Griffin ML, Thase ME, et al. Retention in psychosocial treatment of cocaine dependence: predictors and impact on outcome. *Am J Addict*. 2002;11(1):24–40.