

研究課題名

機械学習を用いたヒト PPAR α に結合する PFAS の特徴の解明

共同研究者名

石橋 弘志 (愛媛大学)

平野 将司 (東海大学)

飯田 緑 (九州工業大学)

研究目的

PFAS (ペルフルオロアルキル化合物およびポリフルオロアルキル化合物) は撥水剤や消火剤として広く使用されている。Forever chemicals と呼ばれる通り、自然界や体内では分解されにくく蓄積されやすい。また、PFAS の中には核内受容体 PPAR α に結合して毒性を示ものがある。これまで、PFAS の個々の構造的特徴 (例えば、炭素原子数や官能基数など) と PPAR α への結合のしやすさについては定量的構造活性相関 (QSAR) などを用いた研究がなされているが、どのような構造的特徴を持つ PFAS が PPAR α に結合しやすいかは明らかとなっていない。これは、PFAS の個々の構造的特徴と PPAR α への結合のしやすさにほとんど相関がないためである。このことから、我々は PFAS の複数の構造的特徴が組み合わせられることで PPAR α への結合しやすさを決定していると考えた。しかし、従来の解析手法では、複数の構造的特徴の影響を検出することは難しい。これは従来法が線形関係かつ単一要因を仮定しているからである。そこで、本研究では、PPAR α に結合する PFAS の構造的特徴の組み合わせを明らかにするために説明可能な人工知能技術 (XAI) を応用する。本研究の遂行により、PPAR α に結合しやすい PFAS の構造的特徴が明らかとなり、毒性が高い PFAS をスクリーニングするための迅速で簡便な手法を構築できる。

研究方法

まず、MOE によるドッキングシミュレーションによって、EPA のデータベースに含まれる 6798 種の PFAS の構造情報から 206 個の記述子と PPAR α 結合に関する S-score を算出した。次に、特徴量選択によって記述子の数を削減した。具体的には、ほとんどの PFAS で同じ値をとる記述子を削除し、さらに、相関を持つ記述子を削除した。その後、再帰的特徴量削減によって、記述子を重要性でランク付けした。ランク付けされた上位の記述子を使って、S-score を予測する機械学習モデルを作成した。機械学習モデルとしては、勾配ブースティング決定木の 1 つである XGBoost を用いた。各記述子が S-score にどのような影響を与えるのか明らかにするために XAI 技術の 1 つである SHAP を用いた。

SHAP は、ゲーム理論においてプレイヤーのチームへの貢献度を計算する手法を一般化したものである。

研究成果

特徴量選択を行った結果、206個の記述子を65個まで削減することができた。さらに、その上位の記述子を使って複数の予測モデルを構築したところ、わずか3個の記述子を使ったモデルでも高い精度で S-score を予測できた ($R^2 = 0.72$)。次に、それら3個の記述子の重要度を SHAP で計算すると、b_single が最も重要度が高く、次いで BCUT_PEOE_3、PEOE_VSA_PPOS となった。b_single は単結合数を示す記述子であることから、PPAR α に結合するには PFAS の分子サイズが最も重要であることがわかった。また、PFAS の単結合数が約60、炭素数が約20、分子量が約800のときに、最も PPAR α と結合しやすいことがわかった。

次に、20個のレガシーPFASと13個の新興PFASに注目して解析を行った。PFOA は POPs 条約で製造や使用が禁止されているが、我々の予測モデルは、9個の新興PFASがPFOAよりもPPAR α に結合しやすいと予想した。一般に新興PFASはレガシーPFASよりも健康に与える影響が小さいと考えられているが、これが必ずしも正しくないことが示唆された。

新興PFASの多くはエーテル結合を持つ。SHAPを使うことで、興味深いことに、エーテル結合が多いPFASほどPPAR α に結合しやすいことがわかった。次に、エーテル結合がどのようなメカニズムでPPAR α への結合に寄与しているのか調べたところ、エーテル結合が増えて分子サイズが大きくなることでPPAR α に結合しやすくなっているわけではなく、その他の要因が関係していることが示唆された。おそらく、エーテル結合が増えることで回転自由度が高まり、PFASがPPAR α に結合しやすくなっていると考えられる。

今後の課題

本研究では、MOEによるドッキングシミュレーションとXAI技術によって、PPAR α と結合しやすいPFASの特徴を明らかにした。これにより、毒性が高いPFASを迅速にスクリーニングすることが可能になった。今回はPPAR α を対象を絞ったが、今後は他の受容体にも本手法を適用し、リスクの高いPFASをスクリーニングできるようにしたい。