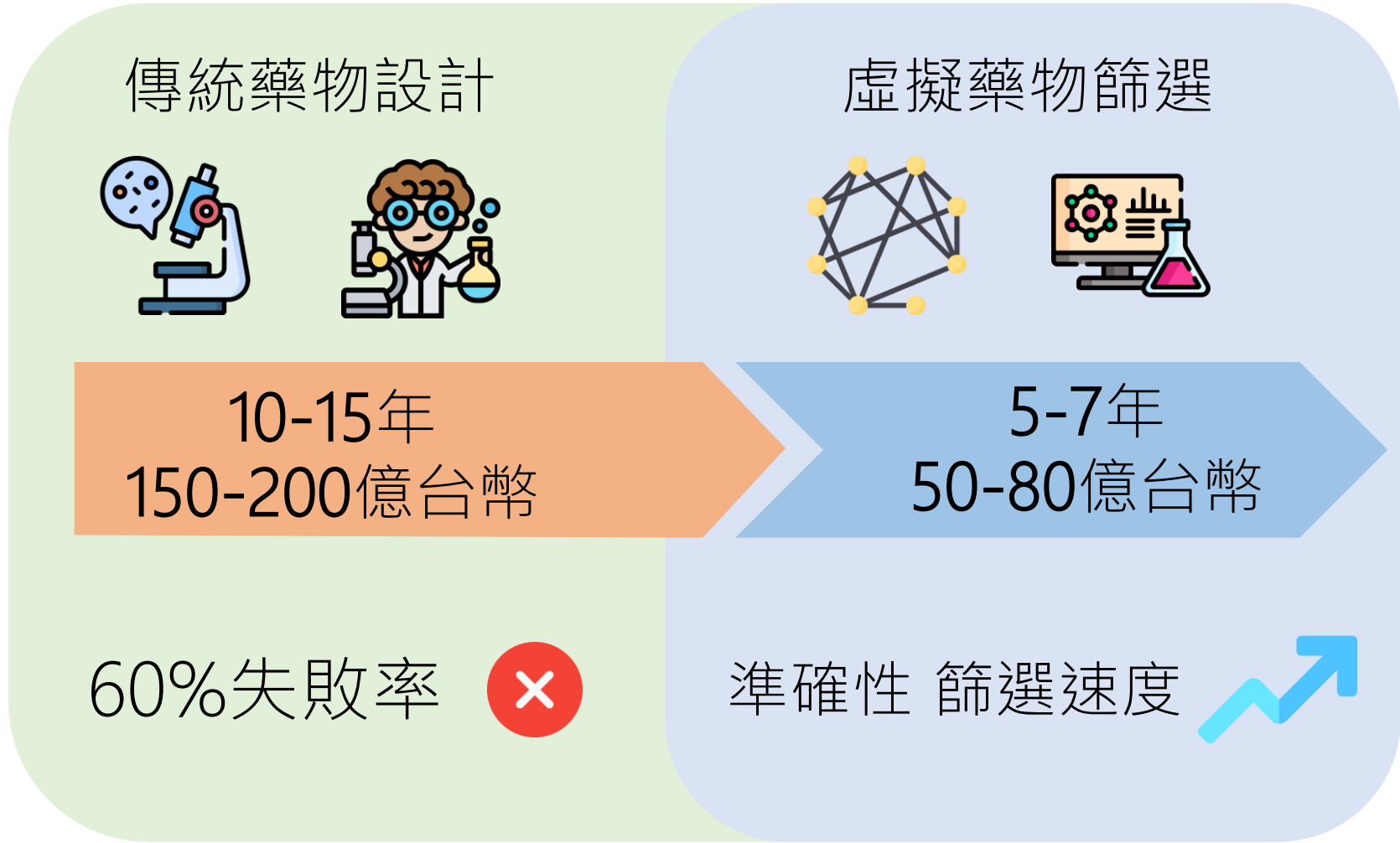
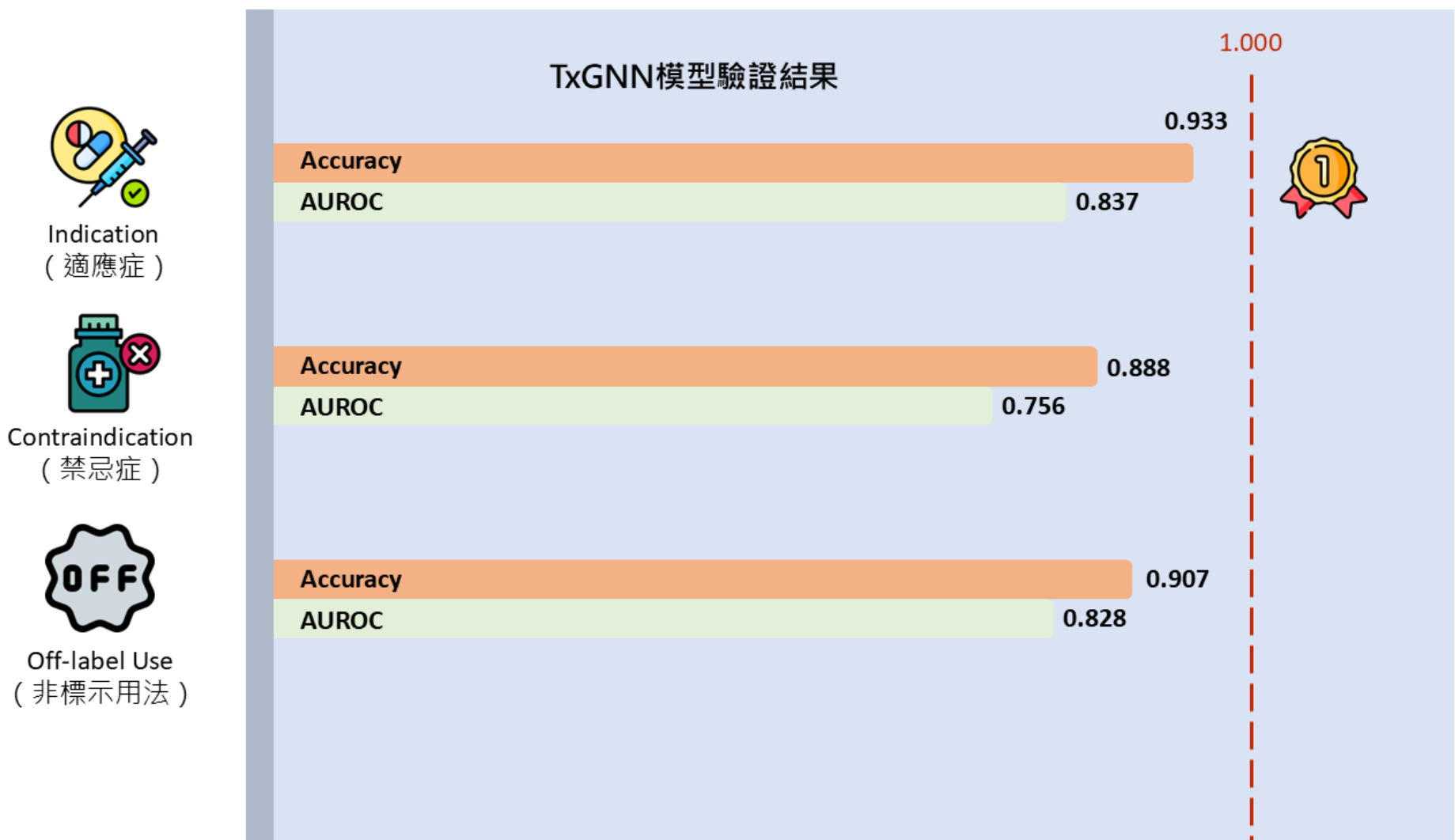


整合知識圖譜進行圖像檢索增強生成技術於虛擬藥物篩選平台



藥物設計耗時費力，在50,000個化合物中，僅有十二分之一能進入人體臨床試驗並最終上市。

	名稱	說明
方法	虛擬藥物篩選	處理速度快、成本低等優勢且能更清晰地理解分子機制
工具	知識圖譜	將複雜關係轉化為節點和邊
	TxGNN幾何深度學習模型	訓練分析疾病與藥物之間的關係



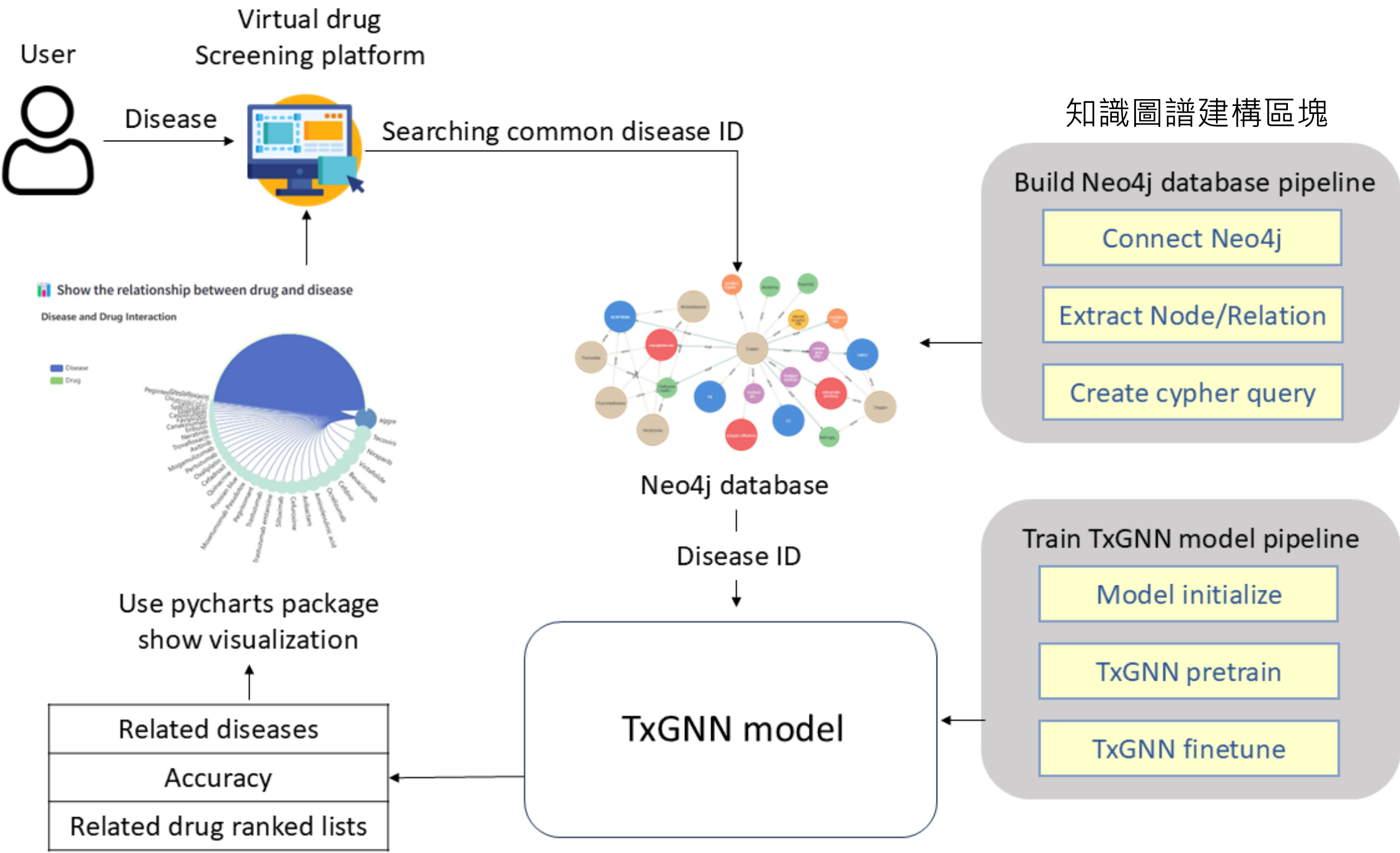
TxGNN在面對Contraindication、Indication和 Off-label Use三種關係時，

Indication 取得最高的 Accuracy為 0.933，Off-label Use最低為 0.888，

表示模型在預測 Indication 關係的準確率最佳，整體分類表現最優。Indication 的 AUROC比另外兩種關係分數高，

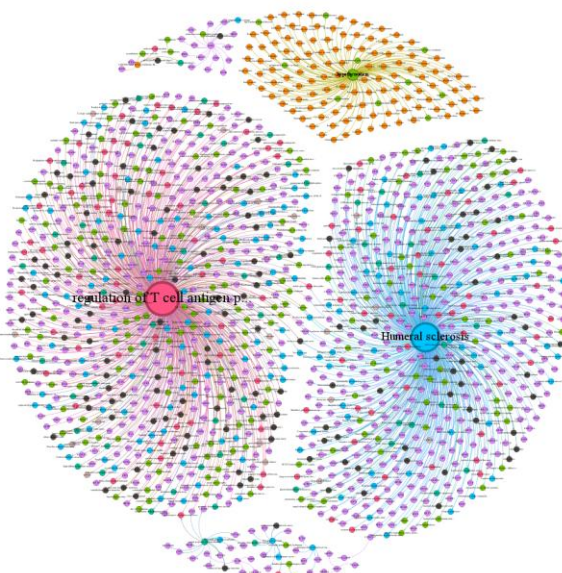
而AUROC越高，代表模型的區分能力 越好，表示模型在區分Indication關係的正負樣本時能力最強，Contraindication則最 弱。

AUPRC主要用於衡量高度不平衡數據中的預測能力，模型在面對三種關係的分數都偏低，顯示模型在識別少數類別時容易出錯。



檢索與預測區塊則為迭代流程，在使用者每次向虛擬藥物篩選平台提交查詢時啟動。系統結合TxGNN預測結果並透過PyCharts套件將藥物排序列表與疾病關聯進行視覺化呈現。最終，透過Streamlit套件將生成結果呈現於直觀的 虛擬藥物篩選介面，供使用者參考。

TxGNN模型訓練區塊同樣為一次性執行流程，利用PrimeKG資料集進行訓練。TxGNN模型流程包括模型初始化訓練與預測資料集分割、模型預訓練模型微調、預測準確度驗證以及訓練效能評估。這些步驟確保模型能夠有效學習知識圖譜中的結構化資訊。



Virtual drug screening platform

Extract relation disease and drug

actinic keratosis (disease)

Select disease: actinic keratosis (disease)

Confirm

Search Result

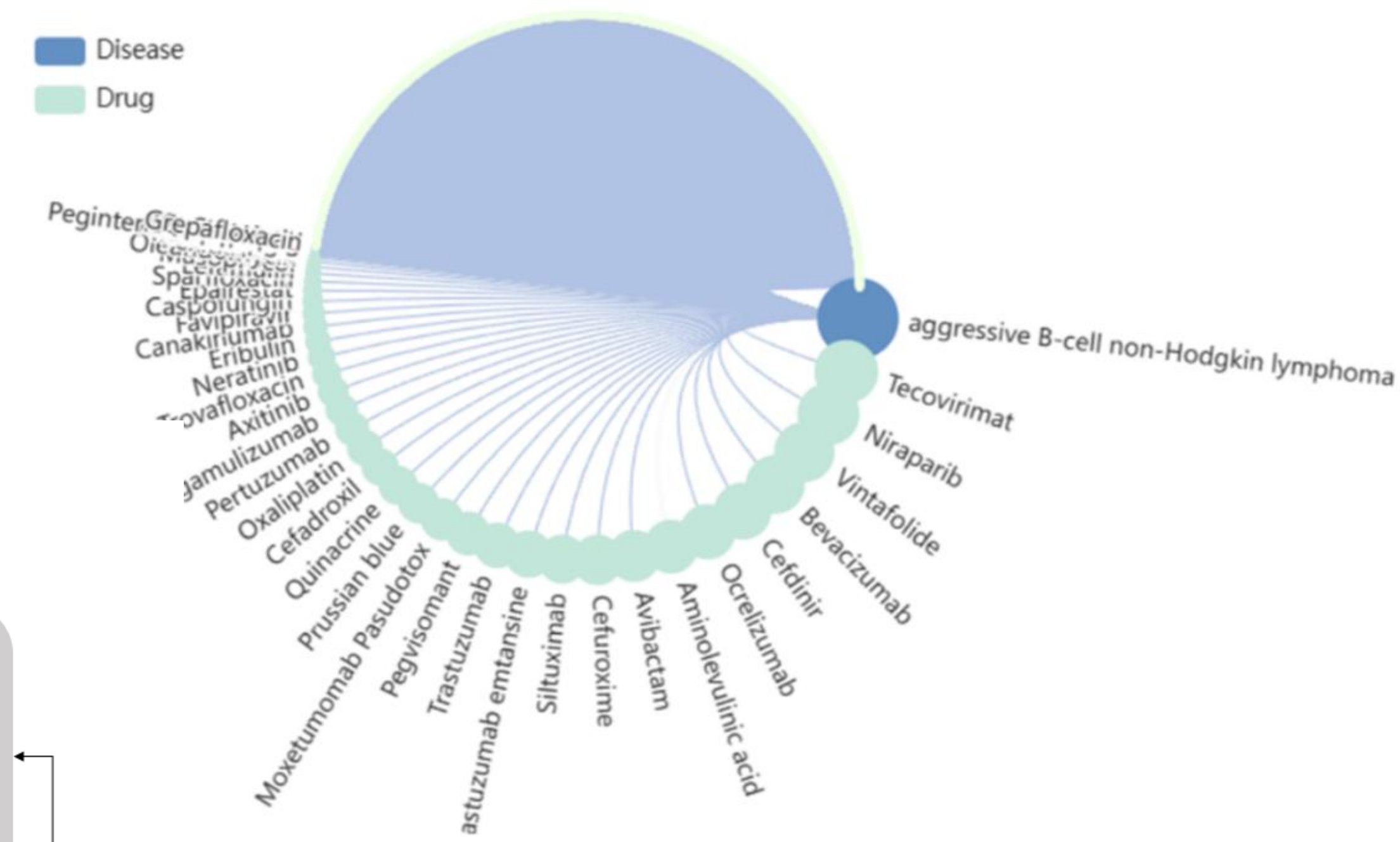
	ID	Name	Accuracy	AUROC	AUPRC
17595.0	17595.0	aggressive B-cell non-Hodgkin lymphoma	0.9062	-1	-1

Ranked list

	aggressive B-cell non-Hodgkin lymphoma
0	Tecovirimat
1	Niraparib
2	Vintafolide
3	Bevacizumab
4	Cefdinir
5	Ocrelizumab
6	Aminolevulinic acid
7	Avibactam
8	Cefuroxime
9	Siltuximab

Show the relationship between drug and disease

Disease and Drug Interaction



知識圖譜建構區塊為一次性執行流程（圖中灰色區塊），始於PrimeKG資料集的導入。Neo4j資料庫建構流程將數據轉化為知識圖譜，包括建立與Neo4j資料庫的連接、從PrimeKG資料集擷取節點與關聯資訊、使用Cypher查詢語法在Neo4j中建立節點與關聯，最後驗證所有節點的成功建立狀態。

此整合方法結合了知識圖譜的結構優勢與幾何神經網路的預測能力，創建出全面性的 虛擬藥物篩選平台，為疾病治療藥物開發提供有效的虛擬篩選工具。

我們希望這項研究能夠找出目前未知的藥物與疾病之間的關係，不僅有助於理解疾病的機制，也為未來的藥物設計提供了新的發展方向。未來研究可結合AlphaFold 深度學習AI模型，預測潛在治療化合物的蛋白質三維結構，並整合分子對接(docking) 技術，通過建立合理的藥效團模型、準確測定或預測標靶蛋白質的分子結構，計算候選化合物與標靶的相互作用。透過這些創新方法和技術，我們期望能加速藥物發現的進程，最終改善患者的治療效果。