

Radiol Oncol 2026; 60(2): 227-243.

doi: 10.2478/raon-2026-0026

Na poti k multimodalni integraciji raka debelega črevesa in kronične ledvične bolezni. Transkriptomsko modeliranje kot okvir za raziskavo SIRIO, imenovano »Prostorska radiomika in transkriptomika za odkrivanje povezave med rakom debelega črevesa in kronično ledvično boleznijo«

Fusco R, Granata G, Belli A, Perna AF, Capasso G, Caraglia M, Pace U, Delrio P, Docimo L, Gambardella C, Lucido FS, Floris M, Locci G, Runfola M, Giannascoli D, Izzo M, Sorgente E, Borriello M, Izzo F, Simeoni M, Petrillo A in sodelavci raziskave SIRIO

Izhodišča. Rak debelega črevesa in danke (*angl. colorectal cancer, CRC*) ter kronična ledvična bolezen (*angl. chronic kidney disease, CKD*) sta glavna dejavnika globalne obolevnosti in umrljivosti zaradi raka. Vse več epidemioloških in genetskih dokazov kaže na biološko verjetno medsebojno vplivanje med ledvično disfunkcijo in nastankom tumorjev v debelem črevesu in danki. Vendar pa javno dostopni nabori podatkovni redko združujejo strukturirane parametre ledvične funkcije z večomskimi podatki o raku, kar omejuje raziskave mehanizmov in napovednih dejavnikov na osi CRC-CKD.

Metode. Sistematično smo pregledali javne vire, vključno z zbirko podatkov za adenokarcinom debelega črevesa (*angl. the cancer genome atlas–colon adenocarcinoma, TCGA-COAD*) in javno dostopnim naborom podatkov o preživetju na podlagi transkriptomskih podatkov, da bi ocenili izvedljivost združevanja kliničnih in genomskih podatkov za odkrivanje bioloških označevalcev. Transkriptomske podatke 62 bolnikov z rakom debelega črevesa smo analizirali z uporabo nenadzorovanega združevanja v skupine, izbire značilnosti na podlagi korelacije ter več nadzorovanih klasifikatorjev za strojno učenje, da bi identificirali genske podpise, povezane s preživetjem brez znakov bolezni.

Rezultati. Projekt TCGA-COAD je potrdil značilno mutacijsko sliko raka debelega črevesa in danke, vendar ni vseboval strukturiranih podatkov o delovanju ledvic. Nasprotno pa je nabor podatkov o preživetju omogočil integrativno modeliranje preživetja brez znakov bolezni. V okviru nenadzorovane analize smo identificirali 3 transkripcijsko različne podskupine. Najvišjo napovedno zmogljivost sta dosegli metodi naključni gozd (*angl. random forest*) in logistična regresija. Analiza pomembnosti značilnosti je kot najpomembnejše napovednike tveganja za ponovitev bolezni izpostavila gene *CYP2E1*, *RAB39A* in *ZBTB3*.

Zaključki. Ugotovitve raziskave razkrivajo kritično vrzel v sedanjih javnih zbirkah podatkov glede integriranih podatkov o raku debelega črevesa in bolezni ledvic ter dokazujejo izvedljivost napovednega modeliranja na podlagi transkriptomike. Ta analiza ponuja računalniški okvir za oblikovanje hipotez, ki bo podprl prihodnje multimodalne raziskave, vključujoče ledvične, molekularne in slikovne parametre pri raku debelega črevesa in bolezni ledvic, kar bo glavni cilj raziskave SIRIO.