

別刷

生命を科学する 明日の医療を切り拓く

実験医学

Experimental Medicine

Vol.40 No.6

2022

4

特集

スパコン・クラウドを 生命科学に使う

ペタバイト時代を生き抜くための
システム整備とデータ活用事例

企画／小笠原理

日本国際賞
受賞者インタビュー

Katalin Karikó 博士, Drew Weissman 博士

mRNAワクチン研究に2人を駆り立てる好奇心

いま知りたい!!

はじまる! がん遺伝子パネル検査データの研究・開発への利活用

羊土社

はじまる! がん遺伝子パネル検査データの研究・開発への利活用

福田博政, 河野隆志 (国立がん研究センター がんゲノム情報管理センター 情報利活用戦略室)

C-CATには、保険診療によるがん遺伝子パネル検査の診療情報とゲノム情報が集積され、2022年1月末までに2万8千件を超える症例が登録されている。この集積データを医学研究や新薬の開発のために利活用していただくための「利活用検索ポータル」のサービスが開始されたので、その特徴や利用方法、申請方法などについて解説する。がん種、遺伝子・変異、薬剤名、年齢・性別等の項目を指定しての症例検索が可能なので、研究・開発にぜひ活用していただきたい。

はじめに

がんゲノム情報管理センター (Center for Cancer Genomics and Advanced Therapeutics, C-CAT) は、日本における良質ながんゲノム医療を実現するために、2018年6月に国立がん研究センター内に設立された^{1)~3)}。2019年6月より、2つのがん遺伝子パネル検査、「OncoGuide NCC オンコパネルシステム」と「FoundationOne CDx がんゲノムプロファイル」が保険適用され、その診療情報と検査結果としてのゲノム情報のC-CATへの登録がはじまった。2021年8月からは、リキッドバイオプシー検査「FoundationOne Liquid CDx がんゲノムプロファイル」が保険適用され、その症例情報もC-CATへの登録対象に加わった。2022年1月末時点でのC-CATへの登録症例数は2万8千件を超えている。2021年秋からは、C-CAT集積データを利活用するためのサービスである「利活用検索ポータル (図1)」の利用申請の受付が開始されたので、本稿ではC-CAT集積データの特徴と利活用検索ポータルでどのような解析ができるのかを中心に解説したい。

C-CATの役割と集積データの特徴

C-CATの主な役割としては、以下の3点があげられる³⁾⁴⁾。①保険診療として行われるがん遺伝子パネル検査を受検した患者一人ひとりのゲノム情報および診療情報を、C-CATのデータベース「C-CATがんゲノム情報レポジトリ」に集約し、わが国のがんゲノム医療・研究のマスターデータベースとして適正に管理する。②ゲノム変化の解釈・臨床的意義付けを行うための「がんゲノム知識データベース (Cancer Knowledge DataBase, CKDB)」を構築し、送られたデータを基に、検出された遺伝子変異の意義や効果の見込まれる薬剤、さらに遺伝子変異に紐づく国内臨床試験の情報を掲載した「C-CAT調査結果」を作成して各病院のエキスパートパネル^{*1}に対して返却することにより、診療を支援する。③C-CATに集積されたデータを、大学等の研究機関や製薬会社等の企業に二次利活用してもらうことで新しい診断法・治療薬等の開発を促進させ、最終的には患者・国民に還元させる。

C-CAT集積データの最大の特徴は、単に疾患名(がん種)と変異のデータペアにとどまらず、薬剤名とその効果、有害事象も含めた治療の情報ははじめとして、転帰情報^{*2}、背景情報^{*3}、検査情報^{*4}など広範囲の診療情報が登録されていることにあり、これらの各診

^{*1} エキスパートパネル：がん遺伝子パネル検査の結果を医学的に検討するために、がんゲノム医療中核拠点病院およびがんゲノム医療拠点病院で開催される多職種専門家からなる会議のこと。がん薬物療法の専門医、遺伝医学の専門医、遺伝カウンセラー、病理医、分子遺伝学やゲノム医療の専門家等が参加し、適切な薬剤選択や生殖細胞系列遺伝子変化への対応などが検討される。

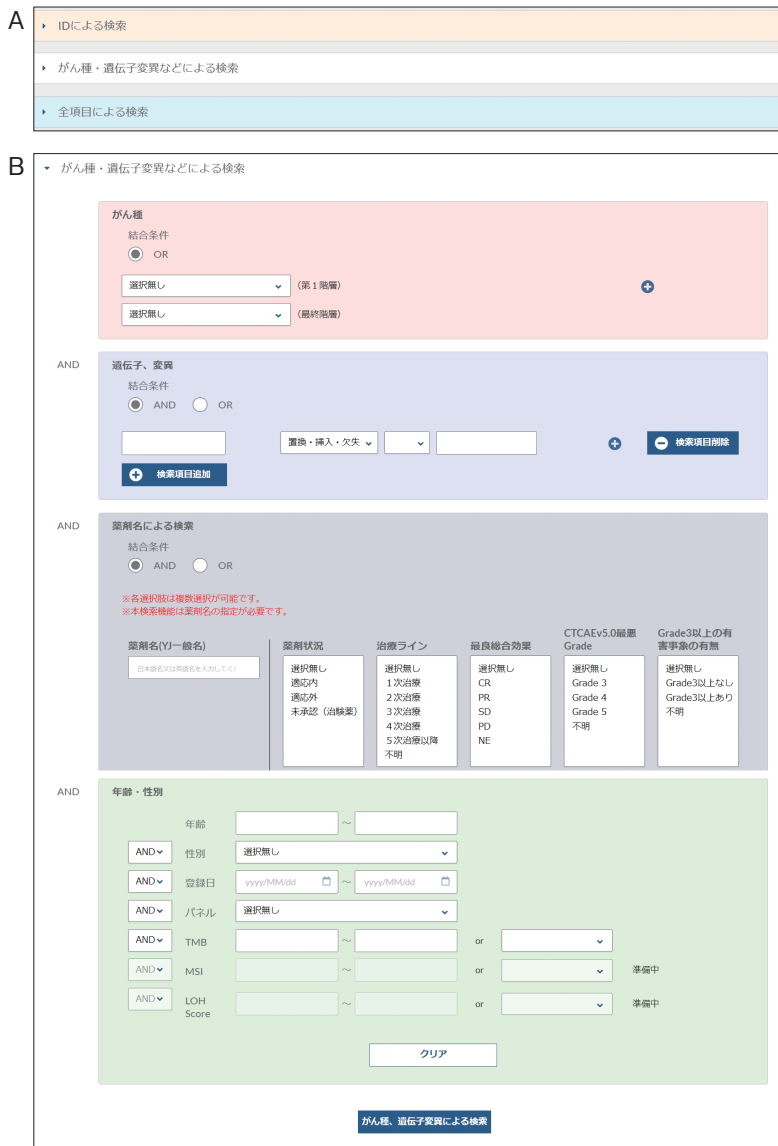


図1 利活用検索ポータルを検索画面

A) 検索トップ画面 (抜粋). 3種類の検索方法から選択できる. B) 「がん種・遺伝子変異などによる検索」を選択した場合の検索画面 (抜粋). がん種, 遺伝子・変異, 薬剤名, 年齢・性別等の4つの検索項目の設定ができる. 4つの検索項目のうち, 単一項目での検索条件の設定も, 複数項目での組み合わせでの検索条件設定も可能である. がん種, 遺伝子・変異, 薬剤名の3つの検索項目では, 「+」ボタンを押すことにより, 新たな入力枠が出現し複数のがん種, 複数の遺伝子・変異, 複数の薬剤での検索条件設定ができる. 薬剤名による検索では, 誌面の都合により一部の条件設定枠を省略している (未省略のイメージは図4参照).

※2 転帰情報: 転帰とは治療における症状の経過や結果を指すが, C-CAT利活用検索ポータルで閲覧できる転帰関連の情報としては, 各治療ライン (レジメン) ごとの完全奏功, 部分奏功等の最良総合効果, 投薬開始・終了日, レジメン終了理由に加えて, 最終生存確認日, 死亡日, 死因がある.

※3 背景情報: 一般的には患者背景情報は, 遺伝要因, 環境要因の他, 年齢, 性別, 既往症など幅広い情報を含むが, C-CAT利活用検索ポータルでは, 年齢, 性別などの基本情報の他, 背景情報として, 病理診断名, 喫煙歴, アルコール多飲, 全身状態, 重複がん, 多発がん, 家族歴の情報が閲覧できる.

※4 検査情報: C-CAT利活用検索ポータルには, がん遺伝子パネル検査前に受検した検査の情報として, EGFR, ALK, HER2等の個別のがん遺伝子の変異・融合・高発現を調べる検査やB型およびC型肝炎ウイルス検査の結果等が登録されている.

療情報と変異との相関解析が可能となっている。収集される臨床項目は、がんゲノム情報レポジトリ臨床情報収集項目一覧表として公開している⁵⁾のでこちらにも必要に応じて参照いただきたい。海外をみても、米臨床腫瘍学会（ASCO）のmCODE⁶⁾や米癌学会（AACR）のGENIE⁷⁾など、学会主導でがんのレポジトリデータベースを臨床試験や基礎研究に役立てようという流れができつつあるが⁸⁾、C-CATの試みはこれらに先行するものであり、わが国の皆保険制度を基盤としている点がユニークでかつ強みとなっている。また、既存の変異データベースは欧米のCaucasianのデータが中心であるので、C-CATデータの解析により日本人に特徴的もしくは東アジア人に特徴的な知見が得られることも期待される。

なお、C-CATへのデータ登録に文書同意された患者さんのデータのみがマスターデータベースである「C-CATがんゲノム情報レポジトリ」へ登録され、そのうち二次利活用への文書同意が得られた症例のデータのみが利活用のためのデータベースに提供され、利活用検索ポータルで利用可能となること⁹⁾、これらに非同意でも保険による検査自体は受検可能であることは明記しておきたい。C-CATがんゲノム情報レポジトリの各症例データは「登録ID」という独自の匿名化IDで管理され、この登録IDと患者氏名やカルテ番号等の個人情報との対応表は各病院にて管理・保管され、C-CATは対応表をもたない。利活用検索ポータルでは、さらに登録IDをハッシュ化した「ハッシュID」により二重匿名化、不可逆化して症例を管理している。また、不正アクセスの監視等、高度な専門的技術により情報セキュリティを高レベルで維持することにも努めている。

利活用検索ポータルで何ができるの？

利活用検索ポータルを利用するとどのような検索ができるかを、実際のスクリーンショットを用いて説明する。検索トップ画面を開くと、「IDによる検索」、「がん種・遺伝子変異などによる検索」、「全項目による検索」の3種の検索方法が指定できるようになっている（図1A）。「IDによる検索」では、自施設症例（がんゲノム医療病院の利用者のみ利用可能な機能）、がん遺伝

子パネル検査の種類、ハッシュID（利活用検索ポータルで用いられる各症例と紐付いた匿名化ID）を指定しての検索ができる。一方、「全項目による検索」では、背景情報や個別の遺伝子検査の情報も含めた任意の項目を指定して組み合わせ検索ができる。ここでは、最も利用機会が多いと思われる「がん種・遺伝子変異などによる検索」を中心に解説したい。

「がん種・遺伝子変異などによる検索」をクリックすると、図1Bに示すような検索条件を指定するための画面が現れる。「がん種」、「遺伝子・変異」、「薬剤名」、「年齢・性別等」の4つの大項目の検索条件が指定でき、この大項目を複数組合わせての検索も可能である。「がん種」では、Onco Tree¹⁰⁾にしたがって、がん種を指定するようになっている。がん種の第1階層をプルダウンメニューから選択すると、第2階層を選択するためのボックスが出現し、第2階層を選択すると第3階層の選択ボックスが出現する。このように順に階層をして、検索したい階層（最終階層）まで指定して、任意のがん種での検索が可能となる。「+」ボタンを押すと別のがん種を指定するためのボックスが追加され、複数がん種の指定が可能となる。

次に「遺伝子・変異」での検索例を示したい。図2Aは、遺伝子にKRASを指定し、置換・挿入・欠失を選択し、G12C変異を指定した場合の画面を示す。「置換・挿入・欠失」の代わりに「遺伝子再構成」を選択すると融合遺伝子の検索ができる。他にも「生殖細胞系列変異」、「コピー数変化」等が選択可能となっている。図2Bには、KRAS遺伝子G12C変異で検索したときの、検索結果を示している。該当症例のサマリーとして、がん種内訳、年代内訳、男女内訳の円グラフおよび変異がコールされた遺伝子のトップ10を示す棒グラフが表示される。グラフ上にカーソルを合わせると、当該箇所件数が表示される（図は腸のがんの領域にカーソルを合わせた例）。グラフの下に該当した症例の件数（254件）と各症例のサマリーテーブルが表示される。病理診断名や薬剤名等、表示が途中で切れている項目はカーソルを合わせることで、隠れている部分を表示できる。

各症例の右端にある「レ」ボタン（図2B赤矢印）を押すと、各症例のレポート（C-CAT調査結果）の変異情報の詳細が表示され（図3A）、「症」ボタン（図2B



図2 利活用検索ポータルでの検索例

A) がん種・遺伝子変異などによる検索の遺伝子・変異の項目で、KRAS 遺伝子/置換・挿入・欠失/G12C 変異での検索条件設定例。 B) KRAS G12C 変異での検索結果 (抜粋)。検索結果の概要として、がん種内訳、年代内訳、男女内訳の円グラフおよび変異がコールされた遺伝子のトップ10を示す棒グラフが表示される。グラフ上にカーソルを合わせると、当該カ所の件数が表示される (図は腸がんの領域にカーソルを合わせた例)。グラフの下に該当症例の件数 (254件) と各症例のサマリーテーブルが表示される。「レ」ボタン、「症」ボタンを押すと、それぞれ各症例のレポートの変異情報、症例情報の詳細が表示される。「レポートCSV」が選択された状態で、右隣の「ダウンロード」ボタンを押すと、チェックで選択した症例のレポートの変異情報がCSV形式でダウンロードできる。「レポートCSV」を「症例CSV」に変更することで、同様に選択症例の症例情報がダウンロードできる。

赤矢印) を押すと、各症例の症例情報の詳細が表示される (図3B)。図3Aの症例では、4つの体細胞変異がコールされ、その一つ目がAPC 遺伝子の Y1376fs*1 (フレームシフト/ナンセンス) 変異であり、次にそのエビデンス情報が示されている。二つ目の変異はKRAS G12Cである (誌面の都合上、途中で画面は切れている)。図3Bでは、表示される症例情報詳細のうち、化学療法の部分のみを抜粋して示している。他にも喫煙歴、家族歴などの背景情報や個別の検査情報なども示される。化学療法レジメン番号 (治療ライン) の行の

いずれかをクリックすると、該当する治療ラインの薬剤名、最良総合効果、有害事象などの詳細情報が表示される。

「レポートCSV」 (図2B 赤矢印) が選択された状態で、右隣の「ダウンロード」ボタンを押すと、チェックで選択した症例のレポート (C-CAT 調査結果) 変異情報がCSV形式でダウンロードできる。「レポートCSV」を「症例CSV」に変更することで、同様に選択症例の症例情報がダウンロードできる。ハッシュIDや検索条件をダウンロードすることもできる。これらの

A

サマリ				
体細胞変異数	生体細胞系列変異数	PMDA承認薬数	国内臨床試験数	PMDA適用外承認薬数
4	0	1	6	0
海外臨床試験数	FDA承認薬数			
0	1			

マーカー	変異番号	変異種類	変異由来	マーカー情報出力区分
APC	1	small_scale_variant	Somatic	True
アミノ酸変化	コピー数変化	サイトバンド	アレル頻度	ToMMoアレル頻度 (表示対象外)
Y1376fs*1		q22.2	0.71 (867/1228)	
海外エビデンスレベル (表示対象外)	米国臨床的意義 (表示対象外)			
概要 (表示対象外)				

※以下のエビデンス情報はC-CAT調査結果作成時のものです。

エビデンス番号	エビデンスタイプ	臨床的意義	遺伝子性疾患名	エビデンスレベル	薬剤	薬剤への到達性
1-1	Predictive	Sensitivity/Response		E	GO07-LK	
1-2	Oncogenic	Likely Oncogenic		F		

マーカー	変異番号	変異種類	変異由来	マーカー情報出力区分
KRAS	2	small_scale_variant	Somatic	True
アミノ酸変化	コピー数変化	サイトバンド	アレル頻度	ToMMoアレル頻度 (表示対象外)
G12C		p12.1	0.51 (753/1471)	
海外エビデンスレベル	米国臨床的意義			

B

症例EP前レジメン情報			
化学療法レジメン番号 (症例EP前)	治療ライン	実施目的	
1	3次治療	根治	
2	2次治療	根治	
3	1次治療	術後補助療法	

症例EP後レジメン情報			
化学療法レジメン番号 (症例EP後)	エキスパートパネル開催日	治療方針	治療ライン
1	2019/09/03	企業治療	2次治療

図3 特定の症例の変異情報詳細と症例情報詳細

A) 特定症例の「レ」ボタン (図2B) を押したときに表示されるレポート (C-CAT調査結果) 変異情報詳細 (抜粋)。**B)** 特定症例の「症」ボタン (図2B) を押したときに表示される症例情報詳細 (化学療法の部分のみ抜粋)。化学療法レジメン番号 (治療ライン) の行のいずれかをクリックすると、薬剤名、最良総合効果、有害事象などの詳細が表示される。

ダウンロードファイルは、研究者の工夫しだいでさまざまな二次解析に利用できると期待される。

図4に、薬剤名にゲムシタピン塩酸塩を入力し (入力アシスト機能がついている)、最良総合効果としてCR (完全奏功) またはPR (部分奏功)、Grade3以上の有害事象なしを選択した場合の条件設定画面を示す。「+」ボタンを押して検索項目を追加することにより別の薬剤との組み合わせ検索ができる。先述のとおり、C-CATデータの“ウリ”は、薬剤名とその効果、有害事象も含めた臨床情報が登録されていることにある。この「薬剤名」の項目と「がん種」、「遺伝子・変異」との組み合わせ検索により、ぜひ利活用検索ポータル

のポテンシャルを引き出していきたい。

利活用検索ポータルを利用するには

利活用検索ポータルを利用するためには、申請書類を提出いただき、情報利活用審査会での承認を経て、C-CATと契約する必要がある。申請には原則として、所属機関等における倫理審査委員会による承認を前提条件としており、倫理審査委員会で承認された研究計画書および承認書の写しを添付する必要がある。ただし、倫理指針の対象外の調査等の場合は、倫理審査委員会による承認は不要とする。利活用検索ポータル利



図4 薬剤名での検索条件設定例

がん種・遺伝子変異などによる検索の薬剤名による検索の項目で、ゲムシタピン塩酸塩、最良総合効果をCRまたはPRとし、Grade3以上の有害事象をなしでの検索条件設定例。「+」ボタンを押して検索項目を追加することにより、別の薬剤との組み合わせ検索ができる。

用料に関しては、データ登録を行っているがんゲノム医療中核拠点病院等からの申請は無償とするほか、アカデミアについては、研究促進の観点より、国としての一定程度の方向性・審査基準のもと採択されている研究に限って無償としているので、ぜひ申請を検討いただきたい。申請手続きや利用料等に関する詳細は、C-CAT公式ホームページの「利活用のページ」¹¹⁾を参照願いたい。

利活用検索ポータル利用申請を検討する段階で、研究対象の症例が何件登録されているかを確認してもらう目的で登録件数検索¹²⁾の機能を公開している。氏名、メールアドレス、所属機関区分、国・地域名の登録だけで、無料利用が可能なので、こちらも利用いただきたい。図1Bに示した利活用検索ポータル検索画面から、「薬剤名」の検索枠だけが除かれた形の検索画面から、「がん種」、「遺伝子・変異」、「年齢・性別等」の各項目を条件設定して検索すると、該当する症例の件数が表示される。利活用検索ポータルと同様のグラフ機能も実装されているので、特定の変異のがん種内訳を出す、特定のがん種で特定の変異を有する症例の年齢分布を出すなど、日々の研究においても気軽に便利使いをしていただきたい。

C-CAT データ利活用の今後の展開

2021年秋より国内機関向けに利活用検索ポータルの申請受付を開始し、2022年1月より契約者の利用がはじまったところである。また、利活用検索ポータルに加えて、利用者がC-CATの提供するクラウド上のプラットフォームにて、がん遺伝子パネル検査のゲノム元データ (FASTQ, BAM等) を含めた解析を行い、その結果をダウンロードできるサービス「利活用クラウド」の準備も進めている。

2022年1月現在保険収載されている3種類のがん遺伝子パネル検査以外にも、造血器腫瘍用のパネル検査を含めて複数のがん遺伝子パネル検査の開発が進行中である。既承認のがん遺伝子パネル検査についても、初回治療時の使用への適応拡大をめざした先進医療での研究が行われている。このようにC-CATへの登録対象となる症例の拡大・増加が予想されている。また、リキッドバイオプシー検査では、その特性を活かして複数回の検査を経時的に実施することが可能であるなど、将来的には同一症例に対して、治療前後、再発時など経時的なゲノム変化の情報が収集される可能性もあると考えている。C-CAT集積データの利用価値は、この流れのなかで、年々増大していくと確信している。

おわりに

本稿をきっかけに、C-CAT利活用検索ポータルの利用者が増え、新たな知見の創出や新規治療薬の開発などを促進し、わが国の医療の向上に貢献するというC-CATの使命が果たされることを祈念する。利用を検討するうえでの疑問点に関しては、C-CAT公式ホームページ「利活用のページ」¹¹⁾に記載の問合せ先メールアドレスまで気軽に照会していただきたい。本稿では詳しく触れなかったが、C-CAT自体についての詳細な説明は、別の総説をお読みいただきたい²⁾¹³⁾。また、C-CATデータ利活用のシステムや運用を改善してよりよいものとするためには、利用者からのフィードバックが不可欠であるので、利用者の方には忌憚らないご意見をお願いしたい。

最後に、本利活用検索ポータルの運用は、C-CATへのデータ登録・利活用にご同意いただいた患者さん、遺伝子パネル検査のデータ・診療情報の登録にご協力いただいたがんゲノム医療中核拠点病院等の方々、遺伝子パネル検査のデータを登録くださった検査企業の方々のご尽力の賜物である。C-CAT一同心より感謝を申し上げるとともに、がんゲノム医療の発展に貢献できるよう、引き続き尽力する所存である。

文献

- 1) 「がんゲノム医療推進コンソーシアム懇談会 報告書」<https://www.mhlw.go.jp/file/05-Shingikai-10901000-Kenkoukyoku-Soumuka/0000169236.pdf> (2017)
- 2) 福田博政 & 河野隆志：腫瘍内科, 25 : 1-6 (2020)
- 3) Mukai Y & Ueno H : Cancer Sci, 112 : 970-977, doi:10.1111/cas.14754 (2021)
- 4) 福田博政 & 河野隆志：がんゲノム医療「遺伝性腫瘍の基礎知識」(平沢見, 編), 304-313, メディカルドゥ (2022)

- 5) 「がんゲノム医療中核拠点病院, がんゲノム情報レポジトリ臨床情報収集項目一覧表」https://www.ncc.go.jp/jp/c_cat/jitsumushya/030/index.html#a3
- 6) 「American Society of Clinical Oncology : mCODE™ : Minimal Common Oncology Data Elements」<https://mcodeinitiative.org/>
- 7) 「America Association for Cancer Reseach : AACR PROJECT GENIE : POWERING PRECISION MEDICINE」<https://www.aacr.org/professionals/research/aacr-project-genie/>
- 8) 大熊裕介, ほか：週刊医学界新聞, 3450 : 4 (2021)
- 9) 中田はる佳, ほか：腫瘍内科, 24 : 183-193 (2019)
- 10) 「Memorial Sloan Kettering Cancer Center : cBioPortal」<http://www.cbioportal.org/oncotree/>
- 11) 「C-CAT公式ホームページ 利活用のページ」https://www.ncc.go.jp/jp/c_cat/use/index.html
- 12) 「C-CAT:登録件数検索」<https://simplesearch-use.c-cat.ncc.go.jp/>
- 13) 福田博政 & 河野隆志：がん分子標的治療, 19 : 19-25 (2021)

参考Webサイト

「がんゲノム医療とがん遺伝子パネル検査」(C-CAT作成) <https://for-patients.c-cat.ncc.go.jp/>

本記事のDOI : 10.18958/6987-00002-0000097-00

Profile

福田博政：1990年東京大学農学部農芸化学科卒業，'95年同大学院農学生命科学研究科博士課程修了。国立がんセンター研究所で、リサーチレジデント、研究員、室長、ユニット長を歴任。その後、2012年よりPMDAでの抗がん薬の薬事審査業務，'15年よりAMEDでのファンディング業務を経て，'18年より現職。趣味は旅行と写真撮影。



Hirokazu Fukuda/Takashi Kohno : Section of Data Science Strategy, Center for Cancer Genomics and Advanced Therapeutics, National Cancer Center

本 pdf は、著者である福田博政先生ご自身によるサーバーへのアップロード、配布の目的にのみご利用いただけます。

以下の行為はご遠慮ください

- ・本 pdf の販売など、直接的な利益を得る目的での利用
- ・本 pdf の一部を抜粋，あるいは改変しての利用
- ・本 pdf の第三者による配布

書籍名：実験医学 Vol.40 No.6

使用頁：909～915 頁

著者名：福田博政，河野隆志

発行社：（株）羊土社

発行年：2022 年

© YODOSHA CO., LTD. 2022