

## がん領域における研究の試料・情報の入手経路

製薬企業では、ヒト由来の試料・情報を利用した研究が数多く行われています。研究を効率的に進めていくためには、生体試料の分譲や医療・生体情報の管理を行う外部の機関や、医療機関の協力が不可欠となります。今回の調査では、現在利用されている試料・情報の入手経路を分析した後に、情報の利用環境に注目して、本邦と海外における環境整備の現状について報告します。

### 1. はじめに

製薬企業では、基礎的な研究にとどまらず、臨床試験や市販後の調査研究を含む医薬品評価のすべての段階で、ヒト由来の試料・情報を利用して、医薬品の有効性や安全性に関する科学的根拠を導き出しています。ヒト由来の試料には、主に血液、体液、組織、細胞、排泄物およびこれらから抽出したDNA等が該当し、情報は、研究対象者の診断および治療を通じて得られた傷病名、投薬内容、検査または測定の結果等、人の健康に関する情報と遺伝情報が該当します<sup>[1]</sup>。

これらの試料・情報は、製薬企業が医薬品に関する研究を行ううえで不可欠なものである一方で、製薬企業が主体的に取得を行うことが困難であるため、生体試料の分譲や医療・生体情報の管理を行う外部の機関や、医療機関の協力が必要となります。製薬企業が質の高い科学的根拠を導き出すためには、研究の目的に合致する試料・情報の入手が必須となるため、これらの機関と密に連携を行っていくことが求められています。

本調査では、製薬企業が関与する研究の試料・情報の供給源を概観するために、近年最も多くの新薬の開発が行われており、試料・情報を提供するための環境整備がほかの領域に先行して進められている悪性腫瘍の領域に注目して調査を行いました。後半では、製薬企業からの利用頻度が高く、情報入手の可否が外部の医療情報提供基盤の整備情報に大きく依存する医療・生体情報の利用環境に注目して、本邦と海外における提供環境の整備状況についても触れます。

### 2. 調査方法

製薬企業における研究に用いられたヒト由来の研究試料・情報の入手経路について、公開されている学術論文をもとに実態を調査しました。悪性腫瘍領域に関連する論文の特定にはWeb of Science (R) クラリベイトのトピック検索を利用しました(検索語<sup>[2]</sup>: “Cancer” OR “Tumor” OR “Oncology” OR “Carcinoma” OR “Sarcoma” OR “Myeloma” OR “Leukemia” OR “Lymphoma”)。上記の検索に該当する論文を2023年に最も多く公表していた悪性腫瘍領域の研究に注力している製薬企業5社(米国拠点3社、英国拠点1社、スイス拠点1社)を特定し、該当する製薬企業の従業員が著者として含まれる2023年に公表された原著論文を今回の調査対象として収集しました。今回の調査では、研究で利用された試料・情報の全体像と分布を把握するために、株化されたヒト細胞や選好を調査するアンケートに伴う属性情報等、容易に入手可能な試料・情報も幅広く含めて分類を行いました。本調査は、対象の企業の国籍や開発パイプラインの状況が結果に影響を与える可能性があり、製薬企業全体の利用実態を定量的に示すものではありません。

### 3. 結果

#### 3-1. 研究に利用される試料・情報

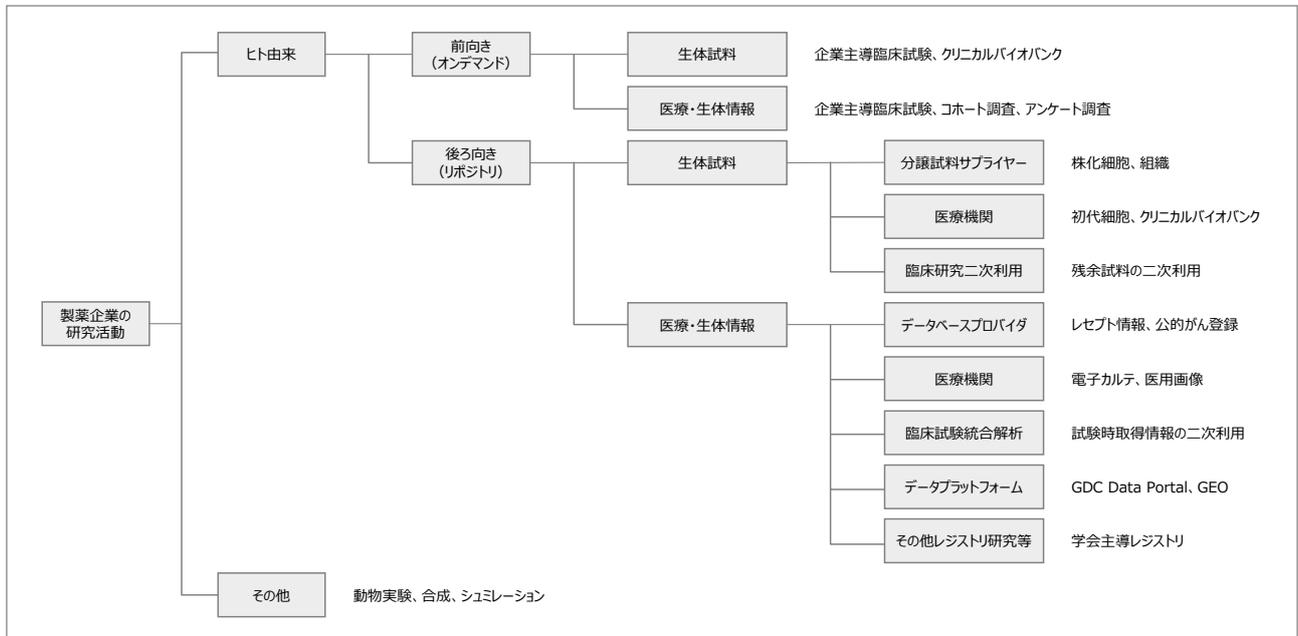
検索の条件に合致した1351報の論文から、内容の確認が困難であった7報を除いた1344報を今回の調査対象としました。今回の調査対象の論文で利用されていたヒト由来の試料・情報を、入手経路という観点で大別すると図1に示すように分類す

[1] 厚生労働省「人を対象とする生命科学・医学系研究に関する倫理指針」  
<https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/hokabunya/kenkyujigyou/i-kenkyu/>

[2] National Institutes of Health, Cancer Classification  
<https://training.seer.cancer.gov/disease/categories/classification.html>

ことができました。入手方法は、必要な試料・情報を研究の目的に沿って前向きに収集する方法と、すでに収集されている試料・情報から、研究の目的に合致するものを選択して利用する方法に大きく分類できます。

図1 研究に利用されるヒト由来の試料・情報の入手先



出所：医薬産業政策研究所にて作成

注：各情報源の枠の右には、各情報源で想定される情報、情報源、利用方法などの一例を示す。

今回の調査対象の論文で利用された試料・情報に基づき作成した分類であり、すべての研究を網羅するものではない。

オンデマンドで前向きに収集した試料・情報を利用する研究は、研究の目的に応じて研究の実施計画を作成し、必要な試料・情報を、必要な量だけ取得できることが利点であり、企業が主導する新薬の臨床試験や患者さんへのアンケート調査、オンデマンド型バイオバンクへの試料提供依頼等が該当します。

すでに収集されている試料・情報を利用する研究は、研究を目的とした収集か否かは問わず、研究立案時点ですでに利用可能な試料・情報を用いることで、試料・情報取得の時間や労力を削減できることが利点であり、サプライヤーから購入した細胞や検体を利用して行う研究や、医療機関の電子カルテの情報を利用する研究等が該当します。

今回の調査対象となっている論文が利用した試料・情報の入手経路の分布を表1に示します。複数の分類の試料・情報を利用して行われた研究は、利用されたすべての分類でカウントしました。医薬品の臨床試験では、有効性評価や有害事象の記録等の医療情報と、血液等の生体試料を同一の試験内で取得するケースが多く、また、生体試料を利用した研究では、論文の記載から、試料取得のための前向きな臨床研究と残余試料を利用した研究の判別が不可能な論文が多数存在したため、表1では、図1に示した分類を一部変更しています。

表1 研究に利用された試料・情報の入手先の分類

情報の種類	論文数 (%) (N = 1,344)
前向き臨床試験・コホート研究・横断研究	499 (37.1%)
生体試料	
サプライヤーの利用 (細胞のみ)	179 (13.3%)
サプライヤーの利用 (その他試料)	27 (2.0%)
医療機関における研究用生体試料の取得 (前向き・後ろ向き)	96 (7.1%)
臨床試験時取得試料の二次利用	30 (2.2%)
医療・生体情報	
データベースプロバイダの利用	146 (10.9%)
医療機関において保管されている情報の利用	61 (4.5%)
臨床試験時取得情報の二次利用	97 (7.2%)
データポータル・データプラットフォームの利用	80 (6.0%)
その他学会主導のレジストリ等の利用	9 (0.7%)
生体試料・情報を利用しない研究	246 (18.3%)

出所：悪性腫瘍領域の研究に注力している製薬企業5社の原著論文をもとに作成

注：複数の情報源に由来する試料・情報を利用して行われた研究は、含まれるすべての分類で1カウントとして集計した。

生体試料のサプライヤー：健康ドナーや患者由来の各種生体試料を研究利用のために集約、提供を行う組織

データベースプロバイダ：医療機関、保険請求、公的調査等の情報を研究利用のために集約、提供を行う組織

データポータル・データプラットフォーム：広く研究に利用するために、ウェブ等で情報を公開するための環境

悪性腫瘍の領域で論文化された、製薬企業が関与するヒト由来の試料・情報を利用した研究において、最も多く利用されていた入手経路は、前向きの臨床試験やコホート研究でした。研究目的に沿って入手された試料・情報から得られた研究結果の質の高さと、臨床試験の透明性を高めるための活動として、研究結果の学術誌への投稿が推奨されていることがこの結果に関連していると考えられます。

続いて、生体試料を利用した研究では、生体試料の分譲を行うサプライヤーからの試料の入手が多い結果となりましたが、容易に入手可能な株化細胞の利用が大半であり、細胞を除く試料に限定すると、医療機関で採取された試料を直接利用した研究がより多く確認されました。医療情報を利用した研究では、医療情報を管理するプロバイダから情報を入手して実施された研究が最も多いという結果でした。

### 3-2. 各入手経路の特徴

試料・情報の入手経路の各分類の中で、今回の調査対象の研究において、高頻度で利用されていた試料・情報の内容と特徴を示します。

#### 3-2-1. 前向き臨床試験・コホート研究・横断研究

医薬品の介入を伴う臨床試験の結果を報告する論文が大半を占めました。その他には、医薬品の製造販売後の調査結果の報告や、安全性監視を目的としたコホート研究、治療の嗜好性等のアンケートを利用した研究等が含まれました。

製薬企業が主導する臨床試験やコホート研究では、研究に参加するすべての被験者から、取得した情報の利用に関する同意を取得するため、情報のアクセス性や研究で必要となる情報の充足度は高くなります。一方で、情報の取得には、人的リソースや時間を要することが多く、長期の追跡や大量の患者情報を収集する必要がある場合には、相応の負担が必要となります。

#### 3-2-2. 生体試料のサプライヤー経由の入手

生体試料の中でも細胞とその他の組織では研究倫理の観点で入手難易度が大きく異なるため、それぞれ分けて整理を行います。細胞を利用した研究では、分譲された株化細胞を利用した研究が大半でした。中でも、米国の非営利団体が試料の提供元となっている研究が大部分を占めました。その他にも日本の細胞バンクや米国の民間のサプライヤーから入手した

研究も存在しました。細胞以外の試料を利用した研究では、組織切片スライドを利用した研究が多く、米国の民間のサプライヤーから入手した研究が大半を占めました。

細胞の入手はサプライヤーから購入することで利用可能であり、試料へのアクセス性は高く、入手時の障害はほとんどありません。本邦における、人を対象とする生命科学・医学系研究に関する倫理指針においても、「既に学術的な価値が定まり、研究用として広く利用され、かつ、一般に入手可能な試料・情報」は指針の対象外とされており、倫理的な観点でも利用への障害は少なくなります。その他の生体試料では、試料の取得時にすべての患者さんから、取得した試料を研究に利用することに同意が取得されることが一般的であり、倫理審査で承認を得ることができれば、試料へのアクセス性は比較的高くなります。

### 3-2-3. 医療機関における研究用生体試料の取得

医療機関における治療期間に採取された腫瘍の生検や血液を利用する研究が多く、医療機関との共同研究や医療機関に併設されるクリニカルバイオバンクやバイオリポジトリ等からの試料の入手が大部分を占めました。医療機関に保管された手術や検査時の残余試料を利用した研究と、研究の目的に応じて前向きに採取を行った研究の双方が存在しました。

これらの生体試料も、試料の取得時にすべての患者さんから、取得した試料を研究に利用することに同意が取得されることが一般的ですが、営利組織である製薬企業での利用が不可能な場合が存在します。前向きに採取を行う場合は、試料採取を目的とした臨床研究となるため、介入を伴う臨床試験と同様に、試料の取得に係る人的リソースや金銭的な負担が大きくなり、すべての試料を入手するまでに時間を要する場合があります。

### 3-2-4. 臨床試験時取得試料の二次利用

臨床試験実施時に採取された残余試料からDNAやRNAを採取し、臨床試験で取得された患者さんの情報と、配列情報や遺伝子発現との関連性を探索する研究が大半を占めました。

これらの生体試料も、試料の取得時にすべての被験者から、取得した試料を研究に利用することに同意が取得されることが一般的ですが、二次利用を行う同意を取得しているか否かが試料へのアクセス性に影響します。

### 3-2-5. データベースプロバイダ経由の入手

民間企業が提供するレセプトや電子カルテに由来するデータベースを利用した研究と、公的機関が提供するがん登録を利用した研究が大半を占めました。民間企業が提供する医療情報の利用では、悪性腫瘍の領域に限定した疾患特異的な情報源(Flatiron、US Oncology Network)だけでなく、疾患横断的な情報源(Optum、Medical Data Vision)から抽出された情報の利用も多く確認されました。公的な患者登録の情報では、米国のがん統計を提供するSEER(The Surveillance, Epidemiology, and End Results)が最も多く利用されており、米国、英国、スウェーデン、デンマークのがん登録も一部の研究で利用されていました。

これらの情報は、プロバイダへの利用料の支払いやアクセスを申請することで利用可能であるため、利用要件を満たせば情報へのアクセス性は高くなります。これらの情報は、各国の規制に基づき個人を特定可能な情報を削除することにより、個人情報保護されているものが多く、大量の患者さんの情報を研究に利用することが容易となることが利点として挙げられます。一方、個人情報を保護する過程で、日付情報や遺伝情報等の研究に必要な情報も削除され情報の充足度が低下する場合があります。

### 3-2-6. 医療機関において保管されている情報の利用

治療期間に取得された、医療機関において保管されている電子カルテの情報を利用した研究が大半を占めました。製薬企業の助成により実施された共同研究と、医療機関の著者が筆頭著者になり製薬企業からの著者が共著者として参加している研究が多く存在しました。病理検査の結果やMRI等の画像情報を利用した研究も一部存在しました。

医療機関では、入院や通院期間に医療機関において行われた処置や医学的所見を網羅的に記録しており、研究で必要となる情報の大部分の項目を入手できる可能性は高くなります。一方で、これらの個人情報を研究に利用するためには、相応の倫理審査や研究体制を含めた情報管理に関する検討が必要となり、情報の取得に掛かる負担は大きくなります。

### 3-2-7. 臨床試験時取得情報の二次利用

有効性や安全性の評価や、母集団薬物動態解析のために複数の臨床試験の結果を統合して解析した研究が多く存在していました。その他にも、費用対効果評価におけるシミュレーションのパラメータ推定での利用や、外部対照群との有効性の比較のために利用した研究も存在しました。

臨床試験時取得試料の二次利用と同様に、情報の取得時にすべての被験者から、取得した情報を研究に利用することに關する同意が取得されることが一般的ですが、本来の臨床試験の目的以外の二次利用も可能となる同意を取得しているか否かが情報へのアクセス性に影響します。

### 3-2-8. データポータル・プラットフォームの利用

トランスクリプトームの発現量や配列情報を利用した研究が大部分を占め、TCGA (The Cancer Genome Atlas Program) やGEO (Gene Expression Omnibus) を利用して情報を入手した研究が大半を占めました。TCGAではDNAの配列や病理標本の画像情報も提供されており、研究数は多くないものの、それらの情報を利用した研究も一部存在しました。

これらのポータルで提供される情報も、一般的に、情報の取得時にすべての被験者から、取得した情報を研究に利用することに関する同意が取得されます。ポータルやプラットフォームを構築し、研究者への情報提供を積極的に行っており、情報へのアクセス性は高く、すぐにダウンロードが可能な情報も存在しています。TCGAでは即時利用可能な情報と、利用の承認が必要な情報が研究倫理の観点で分けられており、配列情報を利用するには一定の利用審査が求められます。

## 3-3. 医療・生体情報の共有基盤

政策研では、医療・生体情報の利用促進に向けた調査を継続的に行っています。今回は、ヒト由来の情報の入手先の中でも、製薬企業からの利用頻度が高く、情報入手の可否が製薬企業外部からの医療情報提供環境の整備状況に大きく依存する「データベースプロバイダ経由の入手」と「データポータル・プラットフォームの利用」に焦点を当て、日本における整備の状況を深掘りしました。今回の調査においては、データベースプロバイダは、利用者の申請と契約に基づきデータの管理・提供を行う組織、データポータル・プラットフォームは、大半の情報がウェブ経由で広くアクセス可能となるように整備された環境と定義しています。

### 3-3-1. データベースプロバイダ経由の入手

今回の調査において、データベースプロバイダ経由の入手の分類で利用頻度の高かった情報を表2に示します。まず、上位のFlatiron Health、SEERの情報はともに米国の悪性腫瘍の領域に特化した情報となっています。前者は、疾患特異的なElectronic Medical Records (EMR) に由来する情報であり、民間企業から提供がなされています。後者は、国立のがん研究所が管理しているがん登録の情報となっています。それらに続く、Optumは、保険請求の情報とEMR等に由来する情報が民間企業から提供されており、全領域の情報から悪性腫瘍の領域の情報が抽出して利用されています。悪性腫瘍の領域では、多様な疾患や用途で利用できる汎用的な情報と、構造化が容易でない疾患特異的な情報の2つが同程度利用されている状況であり、EMRの研究利用もすでに実現され始めていることがわかります。

表2 データベースプロバイダの内訳

情報の入手先	論文数 (%) (N = 146)	地域	主体	対象疾患	主な情報源
Flatiron Health	28 (19.2%)	米国	民間	疾患特異	EMR
SEER: Surveillance Epidemiology and End Results	11 (7.5%)	米国	国立	疾患特異	がん登録
Optum	10 (6.8%)	米国	民間	全疾患	請求データ
IQVIA PharMetrics Plus	7 (4.8%)	米国	民間	全疾患	請求データ
Medical Data Vision	7 (4.8%)	日本	民間	全疾患	請求データ
US Oncology Network iKnowMed EHR	7 (4.8%)	米国	民間	疾患特異	EMR
ConcertAI	5 (3.4%)	米国	民間	疾患特異	EMR

出所：医薬産業政策研究所にて作成

注：利用頻度が5論文以上あった情報の入手先を示した。

本邦においても、悪性腫瘍における疾患特異的な情報の二次利用に向けた動きは始まっており、前述した米国のFlatironも日本に子会社を設立し、国立がん研究センターのSCRUM-Japanプロジェクトと連携を行い、本邦におけるデータベースの構築を目指すことが2022年に公表されています。公的ながん登録情報の民間事業者からの利用可否についても厚生労働省の厚生科学審議会において議論が進められている状況で[3]、全国医療情報プラットフォームの二次利用基盤においても利用可能となる構想が公表されています[4]。

### 3-3-2. データポータル・プラットフォームの利用

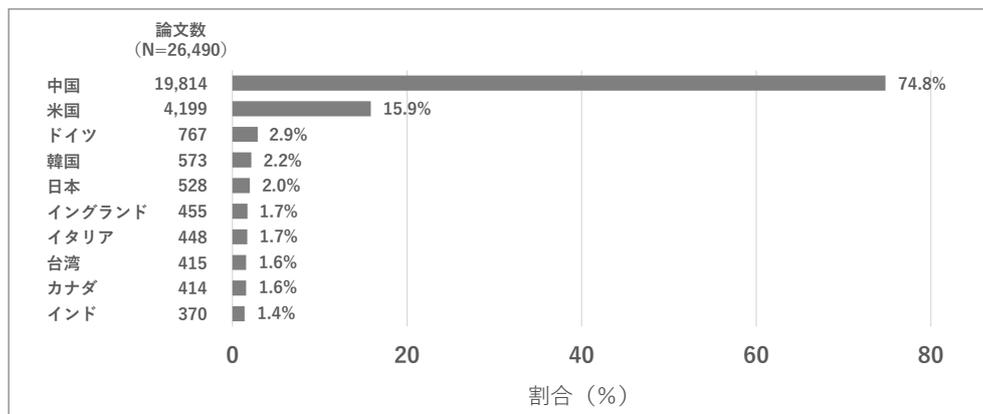
今回の調査において、データポータル・プラットフォーム経由で入手した研究の分類で利用頻度の高かった情報は、TCGAを利用した研究が56%（80報中45報）、GEOに登録された情報を利用した研究が26%（80報中21報）と大半を占め、続いてTARGET (Therapeutically Applicable Research to Generate Effective Treatments) とUK Biobankがそれぞれ5報で続きました。これらは、いずれも米国または英国の機関に管理されている情報源ですが、世界中からの利用が可能となっており、特にTCGAの情報は米国以外の多くの国からも利用がなされています。TCGAに関する記載がある原著論文を、Web of Science (R) クラリベイトで、製薬会社の関与に限らず全論文で確認をすると、中国の著者が関与する割合が米国を上回っており、韓国、日本、台湾等の東アジアの国からの利用も多く見られることがわかります(図2)。このような情報のオープンアクセスが進むことは、健康分野における研究の発展の観点では望ましい動きである一方で、遺伝子の変異情報等は人種によって異なることも報告されており、海外の大規模な情報源のみに依存することへのリスクも厚生労働省の部会で示されています[5]。

[3] 厚生労働省 厚生科学審議会(がん登録部会)「全国がん登録及び院内がん登録に係る課題と対応方針 中間とりまとめ」  
[https://www.mhlw.go.jp/stf/newpage\\_35523.html](https://www.mhlw.go.jp/stf/newpage_35523.html)

[4] 首相官邸 健康・医療戦略推進本部「第11回 健康・医療データ利活用基盤協議会」資料2  
[https://www.kantei.go.jp/jp/singi/kenkouiryou/data\\_rikatsuyou/dai11/gijisidai.html](https://www.kantei.go.jp/jp/singi/kenkouiryou/data_rikatsuyou/dai11/gijisidai.html)

[5] 厚生労働省「第1回がんに関する全ゲノム解析等の推進に関する部会」資料2-1  
[https://www.mhlw.go.jp/stf/newpage\\_07324.html](https://www.mhlw.go.jp/stf/newpage_07324.html)

図2 TCGAに関連する記載がある論文の著者国



出所：Web of Science (R) クラリベイトをもとに作成

注：頻度上位10か国を示した。複数国の情報を利用して実施された研究は、含まれるすべての国で1カウントとして集計した。

本邦においても、遺伝子パネル検査から得られたゲノム情報および診療情報は、国立がん研究センターのがんゲノム情報管理センター(C-CAT)が管理しており、ポータルを経由した利活用が始まっています[6]。その他にも、厚生労働省のもとで全ゲノム解析等実行計画が進められています。がんと難病を対象に、マルチ・オミックスの情報とともに遺伝情報のデータを搭載した情報基盤の構築が行われており、悪性腫瘍の領域で先行して、研究開発に利用できる遺伝情報の情報基盤構築が進んでいる状況です[7]。

#### 4. まとめ

製薬企業が関与するがん研究の試料・情報の供給源の利用頻度の分布を確認し、特に研究への利用の可否が、企業外部の提供体制の整備状況に大きく依存する、医療情報と遺伝情報の現状に関してまとめました。

製薬企業において、研究の目的に沿った試料・情報の取得方法を検討する際に、最も研究目的に合致した試料・情報入手するためには、すべての研究で試料・情報の取得を含めた研究計画を立案し、前向きに取得することが理想となります。しかし、すべての研究で前向きに情報を入手することは、取得にかかる膨大な費用と時間や、大量の試料・情報や希少疾患に関する試料・情報が必要な場合の実施可能性等のさまざまな観点から困難な場合が多いです。このような状況を解決するためにも、研究に必要な時点で、目的を果たすために必要な試料・情報が利用可能となるような環境の構築は、非競争な領域として複数の利用者が協力していくことがイノベーションの促進につながると考えます。

特に、後半で採り上げた、データベースプロバイダにおける情報の整備や情報量・種類の拡大、データポータル・プラットフォーム等による情報提供環境の整備は、製薬企業の研究活動に大きくかかわっており、提供者と利用者が協力し、方向性の検討や情報の整備を行っていくことが望まれます。これらの基盤で取り扱われる医療・生体情報は、国民の健康維持や日本人に向けた医薬品の開発、評価を進めるために非常に重要な役割を果たすため、国が主導する医療分野での、デジタルトランスフォーメーション(DX)推進等の施策の中においても、研究に利用するための情報整備を促進していく必要があると考えます。現在、厚生労働省や国立研究開発法人日本医療研究開発機構(AMED)の事業等において、特に悪性腫瘍の領域でこれらの整備が進められています。このような動きがほかの疾患領域にも波及していくように、迅速に環境整備の検討や利用に向けた制度設計を進めていくことが求められます。

(医薬産業政策研究所 主任研究員 岡田 法大)

[6] 厚生労働省「第4回 がんゲノム医療推進コンソーシアム運営会議」資料1  
[https://www.mhlw.go.jp/stf/newpage\\_17151.html](https://www.mhlw.go.jp/stf/newpage_17151.html)

[7] 厚生労働省「厚生科学審議会科学技術部会全ゲノム解析等の推進に関する専門委員会(第20回)」資料1-1  
[https://www.mhlw.go.jp/stf/newpage\\_38751.html](https://www.mhlw.go.jp/stf/newpage_38751.html)