

# Database Riset Bersumber Data Jaminan Kesehatan Nasional untuk Mendukung Penelitian dan Publikasi Ilmiah di Indonesia

Sanjaya GY<sup>1</sup>, Fuad A<sup>1</sup>, Baros WA<sup>2</sup>, Dhanalvin E<sup>2</sup>, Albab NU<sup>2</sup>

<sup>1</sup>Fakultas Kedokteran, Universitas Gadjah Mada.

<sup>2</sup> Research and Development Group, Badan Penyelenggara Jaminan Sosial (BPJS) Kesehatan

<sup>1</sup>gysanjaya@ugm.ac.id, <sup>2</sup>anisfuad@ugm.ac.id

## ABSTRAK

**Latar Belakang:** Badan Penyelenggara Jaminan Sosial (BPJS) Kesehatan diprediksi akan memiliki data kesehatan terbesar di dunia yang bersumber dari klaim rumah sakit dan pelayanan primer. Selain untuk melakukan monitoring, evaluasi dan pengembangan kebijakan, data tersebut juga berpotensi digunakan untuk penelitian ilmiah. Banyak negara yang sudah memanfaatkan data asuransi kesehatan untuk penelitian, yang kemudian disebut database riset.

**Tujuan:** Penelitian ini bertujuan untuk memahami bagaimana data asuransi kesehatan di dunia dimanfaatkan dan dikelola untuk tujuan penelitian. Dengan demikian, Indonesia dapat mengembangkan database riset yang bersumber data asuransi kesehatan yang dapat bermanfaat bagi lingkungan penelitian.

**Metode:** Literatur review dilakukan untuk menjawab tujuan penelitian tersebut. Pencarian literatur di PubMed dan Google Scholar dilakukan dengan kata kunci tertentu dan dibantu fasilitas find similar article. Untuk mempertajam analisis, penelusuran web penyedia database riset bersumber asuransi kesehatan dilakukan.

**Hasil:** Beberapa negara sudah memanfaatkan data asuransi kesehatan untuk penelitian ilmiah dan publikasi. Data tersebut, yang kemudian disebut database riset dibuat dan dikelola secara khusus oleh lembaga resmi yang memungkinkan diakses oleh peneliti. Sayangnya, data asuransi kesehatan memiliki beberapa kelemahan yang perlu dipahami oleh peneliti. Data sekunder yang dimiliki BPJS Kesehatan sangat berpotensi dikembangkan untuk database riset nasional. Database riset perlu dikelola secara profesional dengan kualitas data yang terus dipantau. Variabel data dapat ditambahkan pada klaim rumah sakit maupun data layanan primer sehingga dapat

dimanfaatkan untuk penelitian kesehatan yang lebih luas.

**Kesimpulan:** Penyediaan database riset bersumber data asuransi kesehatan berpotensi untuk meningkatkan penelitian ilmiah di Indonesia. Volume data, variabel, kualitas dan tatakelola database menjadi pertimbangan penting dalam penyediaan data asuransi untuk penelitian.

**Kata Kunci:** data asuransi kesehatan, database riset, BPJS Kesehatan, literature review

## ABSTRACT

**Background:** Social Health Insurance bureau or BPJS Kesehatan is predicted to have the world's largest health data sourced from claims of hospitals and primary care reports. In addition to monitoring, evaluation and policy development, the data also potentially to be used for scientific research. Many countries have been developed health insurance data for research purpose which is called research database.

**Purpose:** This study aimed to understand how the health insurance data is developed and managed for research. Thus, Indonesia could develop a research database from health insurance data which would be benefit for research environment.

**Method:** The literature review was conducted to answer the research objectives. Literatures were searched in PubMed and Google Scholar with certain keywords and expanded by using Find Similar Article in the systems. To support the analysis, the web page of database research providers were also examined.

**Results:** Some countries already take advantage of health insurance data for scientific research and publication. The research database is created and managed by an official agency that allow interest researchers to access. Unfortunately, health insurance

*data have some limitation that need to be understood by researchers. Secondary data of BPJS Kesehatan are very likely to be developed for a national research database. Database research needs to be managed professionally where the quality of the data should be monitored. Additional variables need to be included in hospital claims data and primary health service report to support wider aspect of health research.*

**Conclusion:** *Availability of research database from health insurance data has the potential to increase*

## PENDAHULUAN

Badan Penyelenggara Jaminan Sosial (BPJS) Kesehatan merupakan lembaga yang ditunjuk sebagai penjamin dalam Jaminan Kesehatan Nasional (JKN). BPJS Kesehatan mulai beroperasi tahun 2014 dengan jumlah peserta 122 juta. Sampai dengan bulan April 2016, jumlah peserta BPJS Kesehatan mencapai 165 juta orang<sup>1</sup> dan akan terus bertambah sampai mencakup seluruh penduduk pada tahun 2019. Dengan demikian, BPJS Kesehatan akan memiliki data kesehatan yang terbesar di dunia. Pada akhir tahun 2015 total data klaim dari rumah sakit melalui INA-CBGs sebanyak 46.850.619 kasus dimana sebagian besar (86,3%) merupakan klaim rawat jalan. Sedangkan jumlah klaim Non INA-CBGs sebanyak 5.569.147 kasus<sup>2</sup>.

Selain jumlahnya yang banyak, variabel data yang dimiliki juga cukup beragam. Data klaim rumah sakit yang menggunakan sistem INA-CBGs memiliki setidaknya 44 variabel yang terdiri dari informasi penyedia layanan kesehatan, informasi pasien dan informasi layanan yang diterima. Sedangkan data layanan primer terdiri lebih dari 38 variabel, tetapi memiliki informasi klinis yang lebih banyak dibandingkan data klaim seperti pemeriksaan fisik, pemeriksaan penunjang dan obat. Dengan kondisi tersebut, data BPJS Kesehatan memiliki potensi yang besar untuk digunakan dalam penelitian<sup>3</sup>.

Sejak tahun 70an, data administratif atau data klaim atau data asuransi kesehatan, sudah dimanfaatkan untuk studi kesehatan<sup>4</sup>. Baru di tahun 2000an jumlah

*scientific research in Indonesia. The volume of data, variables, quality and governance of database is an important consideration in the development of insurance data for research*

**Keywords:** *health insurance data, database research, BPJS Kesehatan, literature review*

penelitian dan publikasi bersumber data asuransi kesehatan berkembang pesat. Di beberapa negara, hasil penelitian bersumber data asuransi kesehatan, telah menghasilkan berbagai publikasi jurnal, konferensi dan laporan ilmiah, baik lokal, nasional dan internasional. Pemanfaatan data tersebut meningkat drastis dari tahun ke tahun, dengan puluhan publikasi per tahunnya<sup>5,6</sup>. Berbagai penelitian klinis, epidemiologi, analisis biaya dan kajian disparitas mutu pelayanan kesehatan telah dilakukan dengan menggunakan data asuransi kesehatan. Hasil penelitian juga telah digunakan untuk pengembangan program kesehatan, monitoring performa pelayanan kesehatan, pengambilan keputusan strategis dan tentunya pembiayaan kesehatan<sup>7</sup>.

Data asuransi kesehatan memiliki sejumlah kelebihan. Data asuransi kesehatan sebagian besar sudah dalam bentuk elektronik. Volume data yang besar mengharuskan pengelolaan data asuransi kesehatan menggunakan sistem elektronik. Dengan demikian akan memudahkan peneliti dalam memilih variabel data yang akan diambil, memilih kasus pada kelompok tertentu, atau mencari kasus-kasus yang jarang ditemukan.

Umumnya data yang keluar untuk digunakan pihak lain sudah melalui tahapan deidentifikasi data, sehingga privasi dan kerahasiaan pasien terjaga. Penggunaan data sekunder mengurangi biaya, khususnya untuk pengumpulan data, yang biasanya harus diambil dari catatan rekam medis fasilitas kesehatan atau pengumpulan data primer melalui wawancara dan penyebaran kuesioner<sup>7</sup>. Data asuransi kesehatan juga

dapat diintegrasikan dengan sumber data lain baik secara agregat berdasarkan kewilayahan atau kelompok umur seperti data survey, maupun data individu seperti data catatan sipil dan statistik vital (kelahiran, kematian, penyebab kematian, pernikahan, perceraian)<sup>8</sup>.

Beberapa negara yang memiliki sistem asuransi kesehatan nasional seperti Taiwan<sup>6</sup> dan Korea<sup>9</sup>, sengaja menyediakan data sekunder tersebut untuk digunakan pihak lain seperti universitas dan lembaga penelitian. Variabel data yang dapat diakses, definisi operasional, standar nilai dan jumlah data yang ada disediakan dalam bentuk informasi yang komprehensif<sup>10</sup>. Jumlah data yang disediakan umumnya data sampling yang diekstraksi sedemikian rupa sehingga representatif terhadap data populasi dan diikuti secara terus menerus. Pertambahan dan pengurangan data dimutakhirkan dalam jangka waktu tertentu, sehingga validitas dan reliabilitas data tetap terjaga dan pada akhirnya berdampak pada kualitas output dari penelitian yang bersumber data asuransi kesehatan<sup>9</sup>.

BPJS Kesehatan sebagai pemilik data asuransi kesehatan bermaksud untuk mengembangkan database riset yang dapat digunakan oleh peneliti dari berbagai lembaga di Indonesia. Bercermin dari negara lain yang lebih dulu mengembangkan database riset asuransi kesehatan, BPJS Kesehatan perlu mengetahui bagaimana database riset dipersiapkan sehingga dapat dimanfaatkan. Selain itu, sebagai lembaga yang relatif baru, BPJS Kesehatan perlu memahami bagaimana pengelolaan database riset bersumber data sekunder. Dengan adanya database riset ini, BPJS Kesehatan dapat membantu meningkatkan penelitian ilmiah di Indonesia dan memberikan dampak terhadap publikasi nasional dan internasional dari Indonesia.

## Metodologi

Literatur review dilakukan untuk menjawab pertanyaan penelitian tersebut. Pencarian literatur melalui PubMed dan Google Scholar menggunakan kata kunci “*health insurance data*” OR “*claim data*” OR “*administrative data*” AND “*research database*”

yang disaring pada judul dan abstrak. Pencarian artikel dibatasi mulai tahun 2000 sampai 2016. Semua artikel yang diperoleh, kemudian diidentifikasi untuk setidaknya memiliki informasi yang berkaitan dengan 1). karakteristik database riset, 2). aksesibilitas data dan 3). deskripsi variabel data. Terdapat 3 artikel yang relevan dengan tujuan penulisan. Fasilitas *find similar article* digunakan pada ketiga artikel tersebut, sehingga didapatkan 5 artikel tambahan yang menunjukkan bagaimana data asuransi kesehatan telah digunakan untuk penelitian di berbagai negara. Full text artikel kemudian didownload dan dianalisis. Analisis diperluas dengan penelusuran web terhadap negara yang memiliki database asuransi kesehatan untuk penelitian atau website organisasi yang mengelola database asuransi kesehatan tersebut.

## HASIL DAN PEMBAHASAN

### 1. Database Riset Asuransi Kesehatan

Data asuransi kesehatan sebagai data sekunder dapat digunakan untuk tujuan penelitian. Beberapa negara menyebutnya sebagai data administrasi atau data klaim yang disiapkan sedemikian rupa menjadi sebuah database riset. Database riset telah tersedia di beberapa negara sejak lama seperti data administratif di provinsi Manitoba dan Saskatchewan, Kanada yang sudah digunakan untuk penelitian sejak lebih dari 30 tahun yang lalu<sup>4</sup>. Taiwan, Korea Selatan dan Prancis telah menggunakan sampel data klaim asuransi kesehatan untuk penelitian yang diambil secara *stratified random sampling* dan dimutakhirkan secara berkala<sup>6,9,11</sup>. Beberapa database riset justru dikumpulkan dari penyedia asuransi kesehatan swasta yang kemudian dimanfaatkan untuk tujuan pengambilan keputusan<sup>12</sup> dan penelitian<sup>13</sup>. Database riset asuransi kesehatan merupakan data yang berasal dari klaim asuransi yang dikumpulkan dan dikelola secara sistematis untuk dapat diakses dan digunakan untuk mendukung pengambilan keputusan maupun penelitian ilmiah.

## 2. Sumber Data pada Database Riset Asuransi Kesehatan

1. Klaim dari rumah sakit
2. Data pelayanan kesehatan di fasilitas kesehatan primer
3. Gabungan dari berbagai penyedia asuransi kesehatan

Sebagian besar sumber data berasal dari klaim yang dilakukan oleh rumah sakit kepada penyedia asuransi kesehatan.

### 3. Dataset

#### a. National Health Insurance Database Taiwan

Taiwan memulai cakupan kesehatan semesta pada tahun 1995. Saat ini, semua penduduk Taiwan telah terlindungi dengan sistem asuransi kesehatan tunggal, dimana sebagian data registrasi dan klaim yang ada dikembangkan untuk tujuan penelitian atau disebut *National Health Insurance Research Database* (NHIRD). NHRID dikelola oleh *National Health Research Institute* (NHRI) Taiwan yang merupakan lembaga independen yang bertujuan untuk meningkatkan penelitian kesehatan di Taiwan. NHRID merupakan salah satu sumber data penelitian yang dimiliki. Peneliti atau lembaga peneliti dapat mengajukan permohonan data NHRID melalui NHRI, namun masih dibatasi pada peneliti dari Taiwan.

Data yang terdapat di NHIRD berasal dari data registrasi dan data klaim asuransi dari semua fasilitas kesehatan di Taiwan. Data sampel diambil (*query*) menjadi beberapa kelompok berdasarkan kriteria tertentu dengan pendekatan *stratified random sampling*. Kelompok data atau dataset dibuat untuk mewakili total populasi dengan kriteria tertentu yang siap digunakan untuk penelitian. Beberapa dataset yang disediakan antara lain:

- Dataset registrasi yang merupakan master data dari registrasi fasilitas kesehatan, registrasi fasilitas spesialis, registrasi tenaga medis,

registrasi fasilitas farmasi dan registrasi kepesertaan

- Sampling Data set yang terdiri dari a). 0.2% sampling data layanan primer yang diambil secara sistematis setiap bulan dan b). 5% sampling data klaim rumah sakit yang diambil secara sistematis setiap bulan
- Data Sampling Kohort (longitudinal), sebanyak 1 juta data individu yang diambil secara random dan diikuti terus menerus.
- Data set spesifik, berdasarkan kriteria inklusi tertentu seperti dataset pelayanan gigi, dataset kanker, dataset penyakit diabetes, dataset psikiatri, dataset penyakit katastropik, kecelakaan lalu lintas dan rehabilitasi medis.
- Teaching demo data set, yang terdiri dari 1.000 data individu yang diambil secara random untuk tujuan pendidikan.

#### b. Clinical Practice Research Datalink Inggris

*Clinical Practice Research Datalink* (CPRD) merupakan kerjasama antara the NHS *National Institute for Health Research* (NIHR), sebuah lembaga independen dan *the Medicines and Healthcare products Regulatory Agency* (MHRA), yang merupakan bagian dari Kementerian Kesehatan Inggris. CPRD bertujuan dalam menyediakan data yang dapat digunakan untuk penelitian. CPRD menyediakan data yang berasal dari catatan aktif pasien yang diambil dari sekitar 650 layanan primer di Inggris. Data layanan primer dikirimkan ke CPRD menggunakan protocol tertentu dengan kesepakatan kelengkapan data socio-demografi dan data klinis yang anonim<sup>5</sup>. Layanan primer di Inggris merupakan bagian dari *National Health System* (NHS) yang pembiayaannya ditanggung oleh pemerintah yang berasal dari pajak.

### c. Sampel Data Klaim Korea Selatan

Korea Selatan telah mencapai cakupan kesehatan semesta sejak tahun 1989 melalui berbagai skema asuransi kesehatan. Baru pada tahun 2000 dikelola secara terpusat oleh *National Health Insurance Service* (NHIS). Sejak itu Korea Selatan menyediakan data sampel representatif terhadap data nasional dan bertujuan untuk digunakan dalam penelitian dan pengambilan keputusan. Secara garis besar terdapat 2 sumber database riset di Korea Selatan, yaitu 1). *the National health Insurance Service – National Sample Cohort* (NHIS-NSC) yang merupakan data kohort dari sampel penduduk tahun 2002-2013, dan 2). Sampel klaim data *Health Insurance and Assessment Service* (HIRA) yang terdiri dari beberapa dataset<sup>10,9</sup>.

Data NHIS-NSC merupakan 2,2% dari total populasi pada tahun 2002 sampai 2013 yang diambil secara *stratified random sampling* yang proporsional terhadap kelompok umur (18 kelompok umur), jenis kelamin, eligibilitas peserta dan level pendapatan (41 level). Terdapat 4 kelompok tabel yang disediakan untuk peneliti, yaitu 1). Data eligibilitas peserta yang merupakan 14 variabel sosio-demografi peserta, 2). Data klaim yang terdiri dari 57 variabel seperti total tagihan, obat yang diberikan, diagnosis dan tindakan medis, 3). Data fasilitas kesehatan yang terdiri dari 10 variabel, serta 4). Data kesehatan umum yang diambil melalui survey dari tahun 2002 sampai 2013 yang terdiri dari pemeriksaan medis umum (*general check up*), gaya hidup dan perilaku kesehatan.

Selain data kohort nasional, terdapat juga data sampel HIRA yang baru dibuat pada tahun 2009. Database HIRA terdiri dari 4 dataset 1). Dataset rawat inap dan rawat jalan, 2). Dataset pasien, 3). Dataset pasien dewasa dan 4). Dataset pasien anak. Masing-masing dataset diambil secara *stratified random sampling* yang proporsional terhadap kelompok umur (16 kelompok umur) dan jenis kelamin (2 kelompok). Masing-masing dataset HIRA terdiri dari 5 kelompok

tabel yaitu 1). Tabel informasi klaim umum yang berisi variabel sosio-demografi pasien, tagihan dan kodifikasi diagnosis, 2). Tabel layanan kesehatan yang terdiri dari variabel tindakan dan penatalaksanaan medis serta obat rawat inap, 3). Tabel diagnosis yang pernah diderita oleh pasien, 4). Tabel obat untuk layanan rawat jalan, serta 5). Tabel informasi penyedia layanan kesehatan.

### d. All-Payer Claims Database dan HCUP Database Amerika Serikat

Di Amerika, asuransi kesehatan dikelola oleh berbagai penyedia asuransi kesehatan yang sifatnya per negara bagian. Penyedia asuransi kesehatan tersebut memiliki data kesehatan yang berasal dari klaim fasilitas kesehatan yang umumnya terdiri dari pelayanan kesehatan, klaim obat, data penyedia layanan kesehatan dan data eligibilitas peserta. Pemerintah Amerika mendorong masing-masing negara bagian untuk mengelola data tersebut dalam bentuk *All-payer Claim Database* (APCD). APCD merupakan kombinasi data yang berasal dari asuransi swasta, *third party administrator* (TPA), asuransi pemerintah, klaim obat dan klaim pelayanan kesehatan gigi yang dikelola masing-masing negara bagian. APCD digunakan untuk berbagai macam tujuan pengambilan keputusan terkait kesehatan, termasuk untuk penelitian<sup>12</sup>.

Pengelolaan APCD juga sangat tergantung dari masing-masing negara bagian. APCD dapat dikelola oleh departemen kesehatan, departemen keuangan atau lembaga independen yang diperkuat dengan peraturan negara bagian. Data tersebut dapat diperoleh dengan permintaan secara online melalui fasilitas yang disediakan oleh masing-masing pengelola APCD di negara bagian<sup>15</sup>.

Selain APCD, secara nasional terdapat database *Healthcare Cost and Utilization Project* (HCUP) yang didanai oleh *Agency for Healthcare Research and Quality* (AHRQ) untuk menyatukan data pelayanan pasien dari negara-negara bagian, asosiasi rumah sakit, organisasi kesehatan swasta, dan pemerintah federal,

sehingga menjadi sumber daya informasi kesehatan nasional. HCUP database memiliki data pelayanan rumah sakit terbesar di Amerika Serikat, yang berasal dari penyedia asuransi kesehatan sejak tahun 1988. Variabel data umumnya terdiri dari identitas unik, jenis kepesertaan asuransi, data sosio-demografi, diagnosis, prosedur medis, obat, penyedia layanan kesehatan, data tagihan (*bill*), tipe fasilitas kesehatan dan tanggal pelayanan. Tersedianya variabel data tersebut memungkinkan penelitian tentang berbagai isu kebijakan kesehatan, termasuk biaya dan kualitas pelayanan kesehatan, pola pelayanan klinis, akses pelayanan kesehatan, dan outcome dari pelayanan kesehatan yang dapat diintegrasikan secara nasional, negara bagian, dan tingkat lokal<sup>16</sup>.

Dataset juga terus berkembang dari sebelumnya 4 dataset menjadi 7 dataset yang tersedia, antara lain 1). the National (Nationwide) Inpatient Sample (NIS) 2). The Kids' Inpatient Database (KID), 3). The Nationwide Emergency Department Sample (NEDS), 4). The Nationwide Readmissions Database (NRD), 5). The State Inpatient Databases (SID), 6). The State Ambulatory Surgery and Services Databases (SASD) dan 7).The State Emergency Department Databases (SEDD)<sup>16</sup>.

#### e. Data Asuransi Kesehatan Kanada

Kanada memiliki skema asuransi yang beragam di setiap provinsinya. Masing-masing provinsi sudah mengupayakan cakupan kesehatan semesta, dimana data administrasi atau klaim telah dikumpulkan secara longitudinal. Beberapa sumber data administrasi tersebut sudah digunakan untuk tujuan penelitian sejak lama<sup>4</sup>, walaupun beberapa provinsi baru mulai menggunakannya<sup>17</sup>.

Pengelolaan database bersumber data administrasi yang juga tergantung dari masing-masing provinsi. Di Provinsi Manitoba dikelola oleh universitas, sedangkan di Saskathewan dikelola oleh lembaga pemerintah. Di Ontario, database administrasi kesehatan dikelola oleh

lembaga independen yang juga mengelola sumber data lain seperti data registrasi fasilitas kesehatan, registrasi kanker dan data survey, sehingga memungkinkan untuk diintegrasikan (*linkage*) satu sama lainnya. Permintaan data untuk tujuan penelitian harus melalui proses seleksi yang ketat. Berbagai macam dataset disediakan seperti dataset asma, penyakit jantung kongestif, transplantasi organ dan resume pasien pulang. Kamus data disediakan secara online sehingga memudahkan peneliti untuk memahami variabel data yang tersedia<sup>18</sup>.

#### f. Statutory Health Insurance Jerman

Asuransi kesehatan di Jerman disediakan oleh 126 penyedia asuransi pemerintah atau *Statutory Health Insurance* (SHI). Berdasarkan peraturan pemerintah, setiap fasilitas kesehatan wajib untuk memberikan data pelayanan ke SHI. Karena banyaknya skema SHI di Jerman, data klaim tersebar di masing-masing SHI, dimana peneliti dapat mengambil satu atau lebih data SHI sebagai sumber data penelitian. Masing-masing SHI memiliki pola pengumpulan data yang berbeda. Bahkan variabel data yang dikumpulkan tidak seragam, sehingga menyulitkan untuk diintegrasikan.

Perkembangan kebijakan kesehatan di Jerman dan penggunaan *Diagnosis Related Group* (DRG) sejak tahun 2005, mendorong semakin banyaknya variabel data asuransi kesehatan yang dikumpulkan dan terstandar dengan kodifikasi yang baku. Secara nasional, kemudian dibentuk *German Institute of Medical Documentation and Information* (DIMDI) yang melakukan pengumpulan data klaim dari semua SHI di Jerman. Data SHI mencakup 86% penduduk Jerman, yang menjadikan DIMDI memiliki data bersumber klaim terbesar yang dapat diakses oleh peneliti di Jerman<sup>19</sup>.

#### g. Database Asuransi Kesehatan Prancis

Prancis memiliki sumber data asuransi kesehatan yang mencakup 65 juta penduduk sejak tahun 2003. Datawarehouse yang dinamakan *the National Health*

*Insurance Inter-Scheme Information Systems* (SNIIR-AM) memiliki dataset sosio-demografi peserta, data layanan rawat jalan, pelayanan medis khusus seperti bedah, obstetri dan ginekologi, psikiatri dan rehabilitasi, data klaim obat, registrasi fasilitas kesehatan, klaim rumah sakit dan data kematian. Semua dataset tersebut dapat saling terhubung menggunakan identitas unik dari individu penerima manfaat asuransi kesehatan. Semua data SNIIR-AM merupakan data layanan kesehatan yang telah diterima oleh peserta<sup>11</sup>.

Selain SNIIR-AM, Prancis juga memiliki sampel data kohort penduduk sebanyak 1/97 jumlah penduduk atau sekitar 660.000 orang, yang secara dinamis diikuti sampai 20 tahun kedepan sejak tahun 2005 atau disebut *permanent beneficiaries sample* (EGB). Sampel data ini termasuk data individu yang belum pernah mendapatkan layanan kesehatan. Jumlah dataset yang dimiliki lebih sedikit dibandingkan SNIIR-AM database<sup>11</sup>.

Kedua sumber data tersebut dapat dimanfaatkan untuk penelitian dengan persyaratan tertentu. Terdapat regulasi khusus yang mengatur bagaimana data dapat diakses, dan bahkan peneliti harus menjalani pelatihan terlebih dahulu sebelum dapat menggunakan data. Konsistensi, kualitas dan ketersediaan data selalu dipantau oleh lembaga yang berbeda yaitu *the health data institute* (IDS). Informasi variabel data atau kamus data tersedia untuk SNIIR-AM database, sedangkan data EGB dapat diakses dan langsung dianalisis secara online.

#### **h. Institutional Health Insurance Database Belanda**

Asuransi kesehatan di Belanda disediakan oleh pihak ketiga. Salah satu penyedia asuransi terbesar adalah *Agis Health Insurance*, dimana data klaim sudah disediakan melalui *Agis Health Database* (AHD) untuk tujuan penelitian. AHD mencakup 1,2 juta populasi yang memiliki data longitudinal sampai setidaknya 10 tahun. Sumber data berasal dari data sosio-demografi pasien, data penyedia layanan kesehatan yang terdiri

dari 2000 layanan primer, 400 apotek dan 20 rumah sakit, serta data layanan kesehatan yang dihubungkan dengan identitas unik individu. Data layanan berasal dari konsultasi dokter umum dan pelayanan spesialis, penggunaan obat, informasi rawat inap, rujukan, fisioterapi, dan keperawatan. Format data diatur secara nasional untuk menjamin validitas data, termasuk penggunaan kodifikasi yang baku<sup>13</sup>.

#### **4. Penjaminan Mutu Database Riset Asuransi Kesehatan**

- Update data secara reguler
- Penggunaan standar kodifikasi baku untuk klaim dan pelaporan

#### **5. Aksesibilitas Database Riset Asuransi Kesehatan**

- Aksesibilitas terbatas
- Fasilitas Online

Berbeda dengan Taiwan, CPRD dapat diakses oleh peneliti dari negara lain. Setidaknya terdapat peneliti dari 22 negara yang telah memanfaatkan data CPRD sampai dengan tahun 2009. Angka ini diprediksi meningkat seiring tren peningkatan jumlah artikel, peneliti dan variasi subjek penelitian yang dilakukan. Secara kuantitas dan kualitas, penelitian bersumber data asuransi kesehatan di Inggris juga menunjukkan peningkatan yang sangat signifikan<sup>5</sup>.

HIRA menyediakan sarana berbasis web untuk mengakomodasi permintaan data untuk tujuan penelitian serta penjelasan tentang variabel data yang dimiliki.<sup>10</sup> HCUP Database dapat diakses publik dengan pengajuan secara online serta menyetujui syarat dan kondisi yang berlaku. Sejumlah biaya diperlukan untuk mendapatkan dataset yang diinginkan.<sup>16</sup>

Permintaan data di Kanada untuk tujuan penelitian harus melalui proses seleksi yang ketat. Kamus data disediakan secara online sehingga memudahkan peneliti untuk memahami variabel data yang tersedia<sup>18</sup>.

Kedua sumber data riset Prancis tersebut dapat dimanfaatkan untuk penelitian dengan persyaratan tertentu. Terdapat regulasi khusus yang mengatur bagaimana data dapat diakses, dan bahkan peneliti harus menjalani pelatihan terlebih dahulu sebelum dapat menggunakan data.<sup>11</sup>

## 6. Publikasi Ilmiah Bersumber Database Riset Asuransi Kesehatan

Menurut sebuah studi dari Chen et al. 2011<sup>6</sup>, jumlah publikasi bersumber NHIRD meningkat drastis dengan rata-rata peningkatan 45,8% per 5 tahun dengan jumlah peneliti yang memanfaatkan NHIRD mencapai 667 orang pada tahun 2009. Bidang penelitian juga semakin bervariasi seperti kesehatan kerja, terapi penyakit, mutu pelayanan, aksesibilitas dan evaluasi pelayanan kesehatan, manajemen pelayanan kesehatan, penelitian kesehatan sosial, ekonomi dan bahkan penelitian ilmu komputer. Dengan demikian, secara kuantitas dan kualitas penelitian bersumber data asuransi kesehatan berkembang pesat di Taiwan.

Pemanfaatan CPRD telah menghasilkan lebih dari 1.500 artikel yang berkaitan dengan keamanan obat, praktek kesehatan terbaik dan *clinical guideline*. Keunggulan dari CPRD karena memuat kohort catatan medis pasien layanan primer yang dihubungkan (*linkage*) dengan beberapa sumber data lain, seperti data layanan rumah sakit (sekunder), data mortalitas dari Badan Statistik Nasional dan data registrasi kanker nasional. Dengan demikian dataset CPRD memungkinkan digunakan untuk penelitian yang lebih luas seperti *pharmacovigilance* (indikasi, pemanfaatan, dan risiko / profil manfaat obat), dukungan untuk penelitian uji klinis (optimasi protokol, kelayakan dan rekrutmen), penelitian kesehatan masyarakat dan menilai outcome pelayanan kesehatan<sup>14</sup>.

Data NHIS-NSC baru dapat diakses pada tahun 2014 dan dalam 8 bulan sudah terdapat 109 penelitian dilakukan<sup>9</sup>.

Berbagai penelitian terkait pelayanan kesehatan, epidemiologi, ekonomi kesehatan dan analisa sumber daya pelayanan kesehatan telah dilakukan dengan memanfaatkan data sekunder bersumber asuransi kesehatan dari SHI. Variabel data seperti obat-obatan, pelayanan rawat inap, data sosio-demografi peserta, diagnosis dan prosedur medis dapat dimanfaatkan bagi peneliti dalam maupun dari luar negara Jerman. Trend publikasi bersumber data asuransi kesehatan di Jerman meningkat, seiring dengan berkembangnya database riset di Jerman<sup>19</sup>.

Secara berkala, dilakukan pertemuan rutin dan membuat publikasi terkait dengan pemanfaatan database riset Prancis agar dapat mendorong pemanfaatan database SNIIR-AM dan EGB oleh peneliti yang tertarik. Berbagai penelitian epidemiologi, farmako-epidemiologi dan ekonomi kesehatan telah dilakukan menggunakan kedua database riset tersebut<sup>11</sup>.

## 7. Tatakelola Database Riset Asuransi Kesehatan

Berdasarkan pengalaman dari berbagai negara diatas, database riset yang bersumber data asuransi kesehatan pada umumnya dikelola oleh lembaga yang independen dan berskala nasional<sup>10,11,16,17,20</sup>. Namun demikian, lembaga ini perlu didukung dengan sumber daya yang mencukupi dan regulasi yang jelas terkait dengan posisinya diantara lembaga-lembaga yang sudah ada. Selain itu, melalui lembaga yang terpusat akan memudahkan koordinasi lintas lembaga, terutama untuk menghubungkan berbagai sumber data yang sudah dikumpulkan oleh lembaga lain.

Dengan adanya lembaga khusus, pengelolaan dan kepemilikan data akan menjadi lebih jelas. Pengelolaan data diantaranya memberikan jaminan kerahasiaan data<sup>11</sup>, melakukan integrasi (*linkage*) dengan sumber data lain<sup>5,8</sup>, menjaga kualitas dan validitas data, mengatur penggunaan data, dan bahkan menyediakan infrastruktur untuk penyimpanan data, aksesibilitas dan analisis data secara mudah<sup>11</sup>. Sangat penting bagi



peneliti untuk mengetahui bagaimana data dikumpulkan menjadi satu datawarehouse yang besar, informasi dataset dan variabel data (metadata) yang tersedia, sehingga dapat menghindari interpretasi yang salah oleh peneliti atau calon peneliti. Beberapa lembaga bahkan menyediakan pelatihan khusus sebelum peneliti dapat memanfaatkan database riset<sup>11</sup>.

Beberapa lembaga seperti di Korea Selatan, Prancis melakukan pemantauan secara khusus terhadap penggunaan data. Pemantauan ini berupa pencatatan yang sistematis terhadap setiap permintaan data oleh peneliti atau lembaga penelitian, melakukan monitoring terhadap output penelitian yang telah dilakukan serta memfasilitasi pertemuan rutin (konferensi) nasional dan penerbitan laporan berkala terkait ketersediaan dataset, kamus data dan pemanfaatan data untuk penelitian<sup>10,11</sup>.

#### 8. Kelamahan Data Asuransi Kesehatan

Data asuransi kesehatan atau data klaim didesain untuk tujuan manajemen. Data ini mengikuti perjalanan pencatatan yang panjang, mulai dari pasien datang ke fasilitas kesehatan, dilakukan pemeriksaan, penatalaksanaan medis, verifikasi dan kodifikasi sampai siap dikirimkan sebagai klaim oleh penyedia layanan kesehatan. Dalam proses tersebut terdapat potensi bias yang dapat terjadi<sup>21</sup>, sehingga perlu diperhatikan oleh peneliti sebagai pertimbangan dari penjaminan kualitas penelitian.

Jumlah kasus dan metode sampling menjadi salah satu kelemahan pertama. Sampling harus memperhatikan representatif terhadap populasi. Namun demikian, sampling juga harus mencerminkan representasi dari variasi kasus yang terjadi, dimana kedua kondisi tersebut perlu dibuktikan secara statistik.

Bagi negara yang belum mencapai *universal health coverage* maka terdapat kemungkinan akan kehilangan sebagian data karena tidak semua penduduk sudah memiliki asuransi kesehatan. Akibatnya tidak mungkin dapat menilai disparitas cakupan pelayanan kesehatan secara objektif<sup>7</sup>. Pasien yang memiliki beberapa

asuransi kesehatan atau cenderung untuk pembayaran *out-of-pocket* akan sulit diikuti karena klaim tidak terintegrasi atau bahkan tidak ada laporan yang masuk dalam database riset<sup>13</sup>.

Banyak yang mengkritisi variabel data asuransi kesehatan sangat minim, dan sangat tergantung bagaimana prosedur klaim di masing-masing negara. Bagi negara yang menggunakan pendekatan *Diagnostic Related Group (DRG)* untuk klaim misalnya, beberapa kondisi klinis tidak dapat diperoleh seperti komplikasi tindakan medis, potensi *medical error*, informasi selama proses perawatan, tindakan dan penunjang medis ringan, diagnosis penyerta pada kasus penyakit kronis, dan obat-obatan yang diberikan selama perawatan pasien. Namun demikian, kondisi ini akan berbeda bagi negara yang telah menerapkan *electronic health records* secara luas<sup>7,13</sup>.

Terakhir berkaitan dengan kualitas data yang diperoleh dari klaim rumah sakit. Seperti yang diketahui tujuan klaim bagi rumah sakit adalah untuk memperoleh nilai reimburs seoptimal mungkin. Tidak dapat dipungkiri, adanya kemungkinan tindakan *fraud* pada klaim asuransi kesehatan dengan melakukan *up coding* terhadap kasus-kasus tertentu. Akibatnya kualitas data dipertanyakan.

#### 9. Database Riset BPJS Kesehatan

BPJS Kesehatan memiliki data yang dikumpulkan secara internal maupun laporan dan klaim dari fasilitas kesehatan. Secara internal, BPJS Kesehatan mengumpulkan data registrasi fasilitas kesehatan, registrasi peserta jaminan, data pengumpulan premi dan data survey primer yang dilakukan. BPJS Kesehatan juga mengumpulkan data fasilitas kesehatan layanan primer melalui aplikasi PCare serta data klaim dari fasilitas kesehatan tingkat lanjutan (rumah sakit) melalui aplikasi INA-CBGs.

Bercermin dari pengalaman beberapa negara lain dan tantangan pengembangan database riset bersumber data asuransi kesehatan, terdapat 4 pertimbangan dalam

pengembangan database riset yang bersumber data BPJS Kesehatan, antara lain:

1. Volume data yang disediakan berupa sampling yang representatif terhadap populasi peserta BPJS Kesehatan. Pemutakhiran data perlu dilakukan terus menerus dengan jangka waktu sesingkat mungkin mengingat populasi peserta BPJS Kesehatan yang terus bertambah dengan target cakupan kesehatan semesta pada tahun 2019.
2. Untuk mendukung variasi penelitian di Indonesia, dataset yang disediakan dibuat beragam yang dapat merepresentasikan penelitian pada kelompok-kelompok subjek tertentu. Contoh dataset yang dapat dipersiapkan antara lain dataset untuk kohort individu, dataset keluarga, dataset anak, dataset tren kunjungan dan dataset penyakit khusus.
3. Variabel data yang disediakan berasal dari data kepesertaan, data registrasi fasilitas kesehatan, data klaim INA-CBGs untuk pelayanan tingkat lanjutan, variabel data P-Care untuk kunjungan layanan primer dan data klaim penggunaan obat. Definisi operasional dan referensi nilai (*value*) dari variabel data tersebut perlu dibuat dan disebarluaskan agar memberikan gambaran bagi calon peneliti terkait potensi penelitian yang dapat dilakukan.
4. Integrasi database riset dengan sumber data lain dapat meningkatkan variasi penelitian kesehatan maupun non kesehatan. Integrasi data dapat dilakukan secara individu atau agregat. Integrasi data individu dilakukan dengan menggunakan identitas unik sehingga dapat dihubungkan dengan data individu lain seperti data sensus, data catatan sipil, registrasi penyakit menular dan tidak menular serta *electronic medical records*. Sedangkan secara agregat, dapat diintegrasikan berdasarkan wilayah atau kelompok usia tertentu, sehingga dapat dihubungkan dengan data agregat lain seperti data sumber daya kesehatan (jumlah tenaga medis,

jumlah tempat tidur rawat inap) dan data sosial ekonomi penduduk<sup>8,22</sup>.

## KESIMPULAN DAN SARAN

Banyak negara sudah memanfaatkan database riset bersumber asuransi kesehatan untuk penelitian klinis, studi mutu dan disparitas pelayanan kesehatan, epidemiologi, analisis biaya dan kesehatan masyarakat. Variabel data yang tersedia sesuai dengan mekanisme persyaratan klaim masing-masing penyedia asuransi kesehatan. Sebagian besar volume data yang tersedia merupakan data sampel yang dikelola sedemikian rupa sehingga representatif terhadap populasi. Beberapa isu kelemahan data klaim perlu dipertimbangkan dalam membuat database riset. Berdasarkan pengalaman tersebut, BPJS Kesehatan dapat menyiapkan database riset dengan metode sampling, yang dibuat dalam dataset yang beragam. Selain itu, database riset BPJS Kesehatan juga dipersiapkan untuk dapat diintegrasikan (*linkage*) dengan sumber data lain yang sudah tersedia.

Database riset perlu dikelola secara nasional agar terjamin aksesibilitas, akuntabilitas dan transparansi pengelolaannya. Untuk mendukung penelitian yang berkualitas, validitas dan kualitas data perlu dipantau agar sesuai dengan kondisi sebenarnya. Selain itu, variabel data yang penting untuk penelitian klinis perlu diperbanyak melalui prosedur klaim yang ditetapkan, sehingga variasi penelitian klinis dan non klinis dapat dimungkinkan.

## KEPUSTAKAAN

1. BPJS Kesehatan. Jumlah Peserta BPJS Kesehatan. 2016. <http://bpjs-kesehatan.go.id/bpjs/>. Accessed May 10, 2016.
2. Grup Penelitian dan Pengembangan BPJS Kesehatan. Estimasi Biaya Menggunakan The Basic Chain Ladder Method. 2016.
3. Kementerian Kesehatan RI. Kamus Data Kesehatan Indonesia. *Kementerian Kesehatan RI*.

2015. <http://idn-hdd.depkes.go.id/>. Accessed May 23, 2016.
4. Tricco AC, Pham B, Rawson NSB. Manitoba and Saskatchewan administrative health care utilization databases are used differently to answer epidemiologic research questions. *J Clin Epidemiol.* 2008;61(2). doi:10.1016/j.jclinepi.2007.03.009.
  5. Chen Y, Wu J, Haschler I, Majeed A, Chen T, Wetter T. Academic impact of a public electronic health database: bibliometric analysis of studies using the general practice research database. *PLoS One.* 2011;6(6).
  6. Chen Y, Yeh H, Wu J, Haschler I, Chen T, Wetter T. Taiwan's National Health Insurance Research Database: administrative health care database as study object in bibliometrics. *Scientometrics.* 2011:365-380.
  7. Ferver K, Burton B, Jesilow P. The Use of Claims Data in Healthcare Research. *Open Public Health J.* 2009;2(1):11-24. doi:10.2174/1874944500902010011.
  8. Bradley CJ, Penberthy L, Devers KJ, Holden DJ. Health services research and data linkages: Issues, methods, and directions for the future. *Health Serv Res.* 2010;45(5 PART 2):1468-1488. doi:10.1111/j.1475-6773.2010.01142.x.
  9. Lee J, Lee JS, Park S-H, Shin SA, Kim K. Cohort Profile: The National Health Insurance Service-National Sample Cohort (NHIS-NSC), South Korea. *Int J Epidemiol.* January 2016:dyv319. doi:10.1093/ije/dyv319.
  10. Kim L, Kim J, Kim S. A guide for the utilization of Health Insurance Review and Assessment Service National Patient Samples. *Epidemiol Health.* 2014;36:1-4.
  11. Moulis G, Lapeyre-Mestre M, Palmaro A, Pugnet G, Montastruc J-L, Sailler L. French health insurance databases: What interest for medical research? *La Rev médecine interne / fondée . par la Société Natl Fr médecine interne.* 2015;36(6):411-417. doi:10.1016/j.revmed.2014.11.009.
  12. Miller BPB, Love D, Sullivan E, Porter J, Costello A. *All-Payer Claims Databases An Overview for Policymakers.*; 2010.
  13. Smeets HM, Wit NJ De, Hoes AW. Routine health insurance data for scientific research : potential and limitations of the Agis Health Database. *J Clin Epidemiol.* 2011;64(4):424-430. doi:10.1016/j.jclinepi.2010.04.023.
  14. Medicine and Healthcare Product Regulatory Agency. Clinical Practice Research Datalink. 2016. <https://www.cprd.com/home/>. Accessed May 23, 2016.
  15. Agency for Healthcare Research and Quality. All-Payer Claims Databases. *Agency Healthc Res Qual.* 2016. <https://ushik.ahrq.gov>. Accessed May 2, 2016.
  16. Healthcare Cost and Utilization Project (HCUP). HCUP Database. *Agency Healthc Res Qual.* 2015. <http://www.hcup-us.ahrq.gov/databases.jsp>. Accessed May 26, 2016.
  17. Ishiguro L, Saskin R, Vermeulen MJ, Yates E, Gunraj N, Victor JC. Increasing Access to Health Administrative Data with ICES Data & Analytic Services. *Healthc Q.* 2016;19(1):7-9.
  18. ICES Data & Analytic Services. Data and Analytic Services. *Inst Clin Eval Sci.* 2016. <https://datadictionary.ices.on.ca/Application s/DataDictionary/Default.aspx>. Accessed May 27, 2016.
  19. Kreis K, Neubauer S, Klorä M, Lange A, Zeidler J. Status and perspectives of claims data analyses in Germany-A systematic review. *Health Policy.* 2016;120(2):213-226. doi:10.1016/j.healthpol.2016.01.007.
  20. National Health Research Institute. National Health Insurance Research Database. *Natl Heal Insur Adm Minist Heal Welf.* 2014. <http://nhird.nhri.org.tw/en/index.html>. Accessed May 23, 2016.

21. Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. *J Clin Epidemiol.* 2005;58(4):323-337. doi:10.1016/j.jclinepi.2004.10.012.
22. European Center for Social Welfare Policy and Research. Good practice on data linkages and performance measurement in relation to access to national health care data systems. 2013;7(4):55.