

THE WESTERN CAPE PROVINCIAL HEALTH DATA CENTRE

Insight conference 23 May 2017, Cape Town

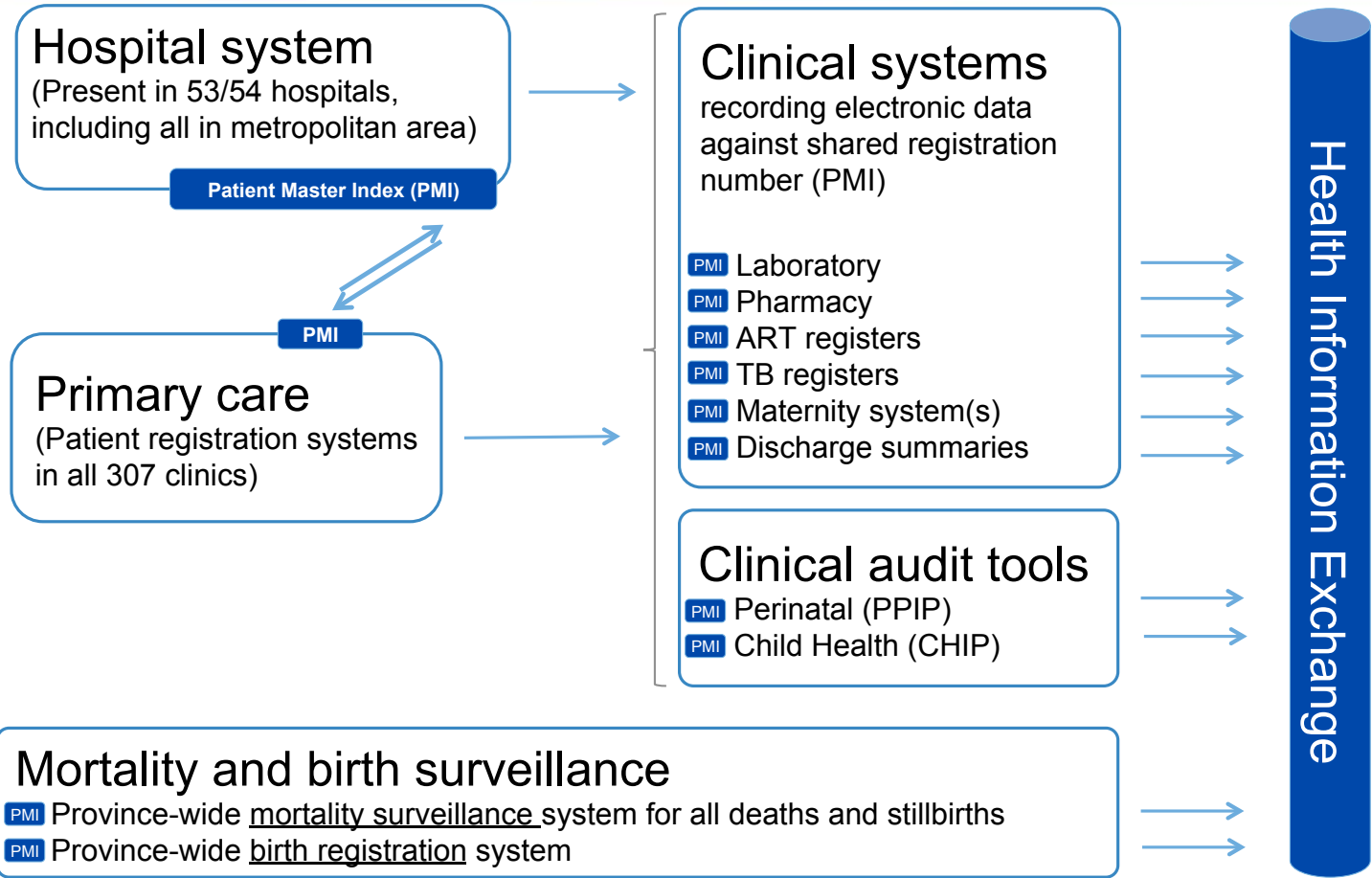
Andrew Boulle

Health Impact Assessment Directorate, Western Cape Department of Health
University of Cape Town, School of Public Health and Family Medicine

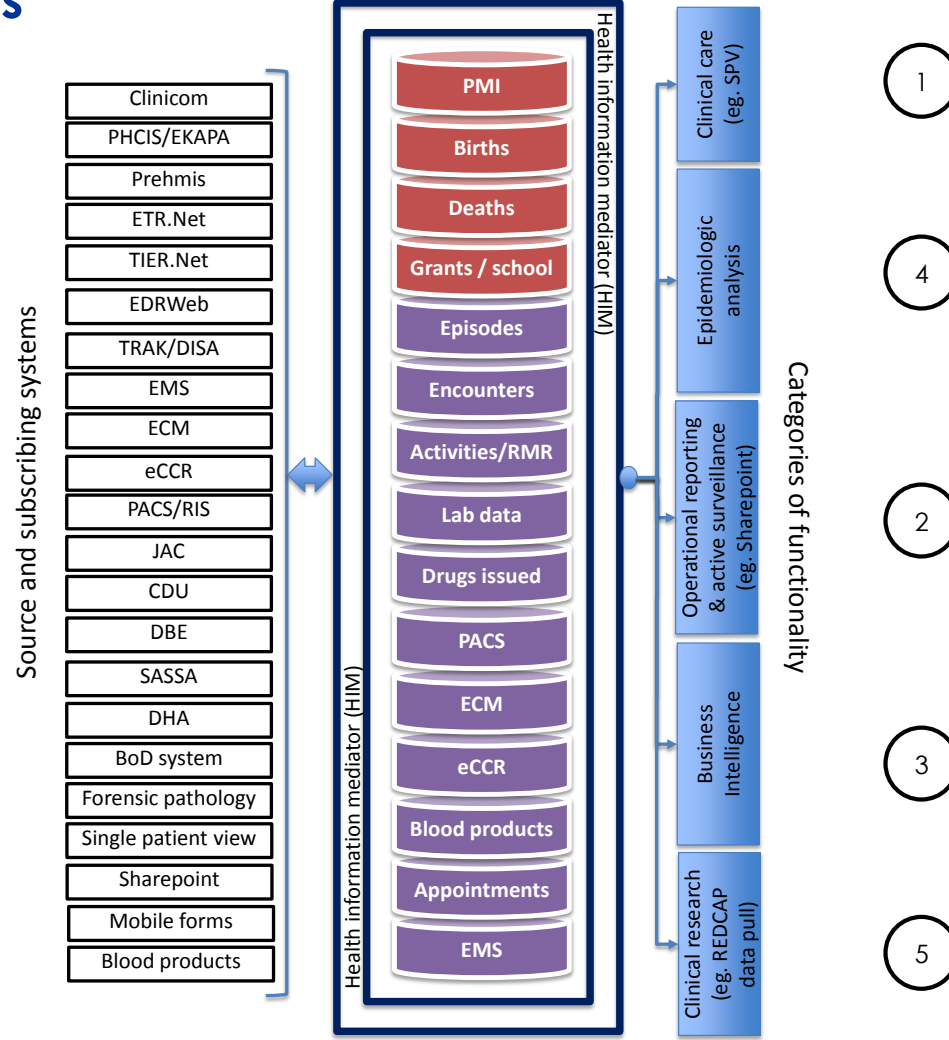
Outline

- Introduction, background, architecture
- Data curation
- Data beneficiation and enrichment: encounters, episodes and cascades
- Reporting
- Data governance, sharing and national alignment

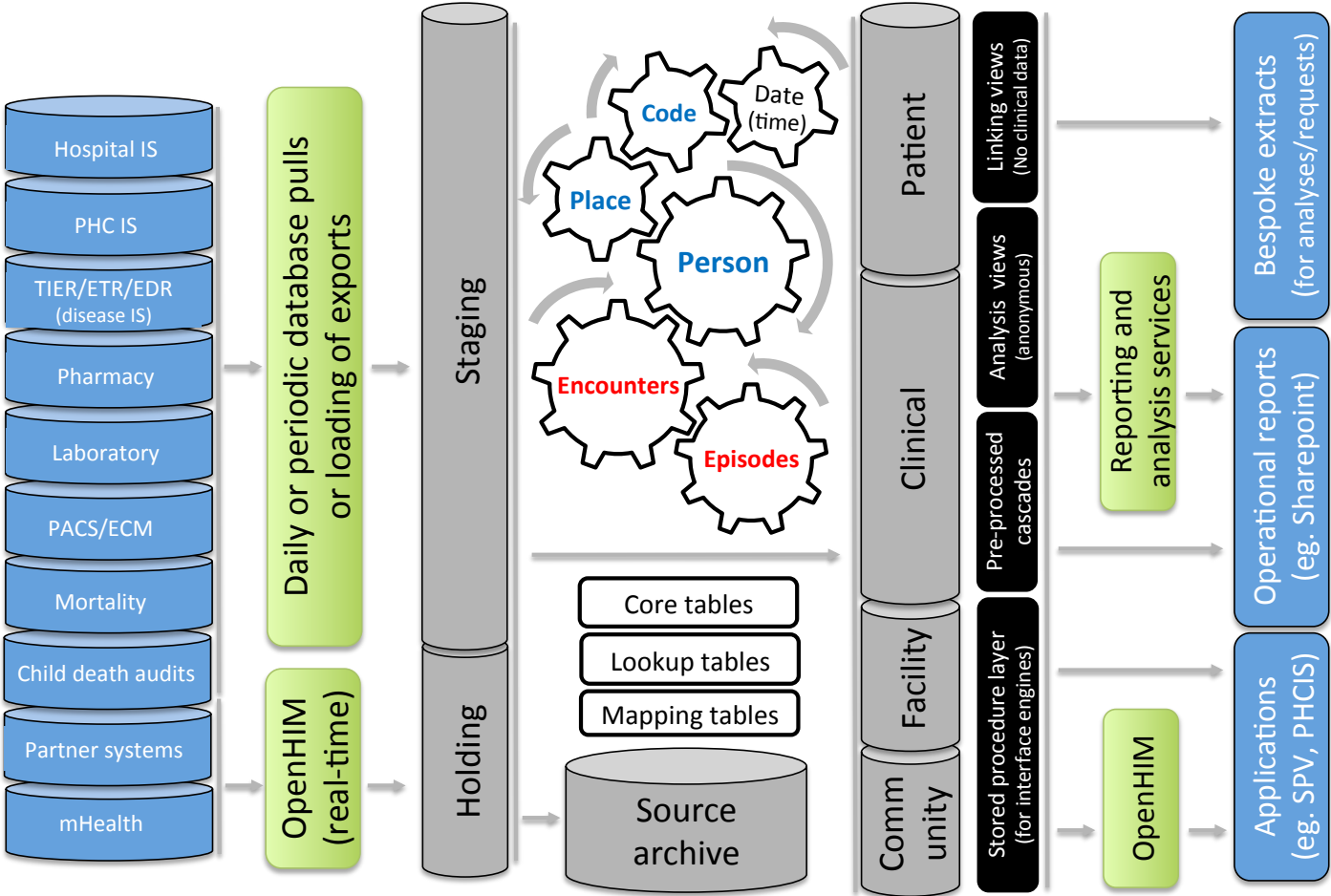
Linkage of data on a unique identifier



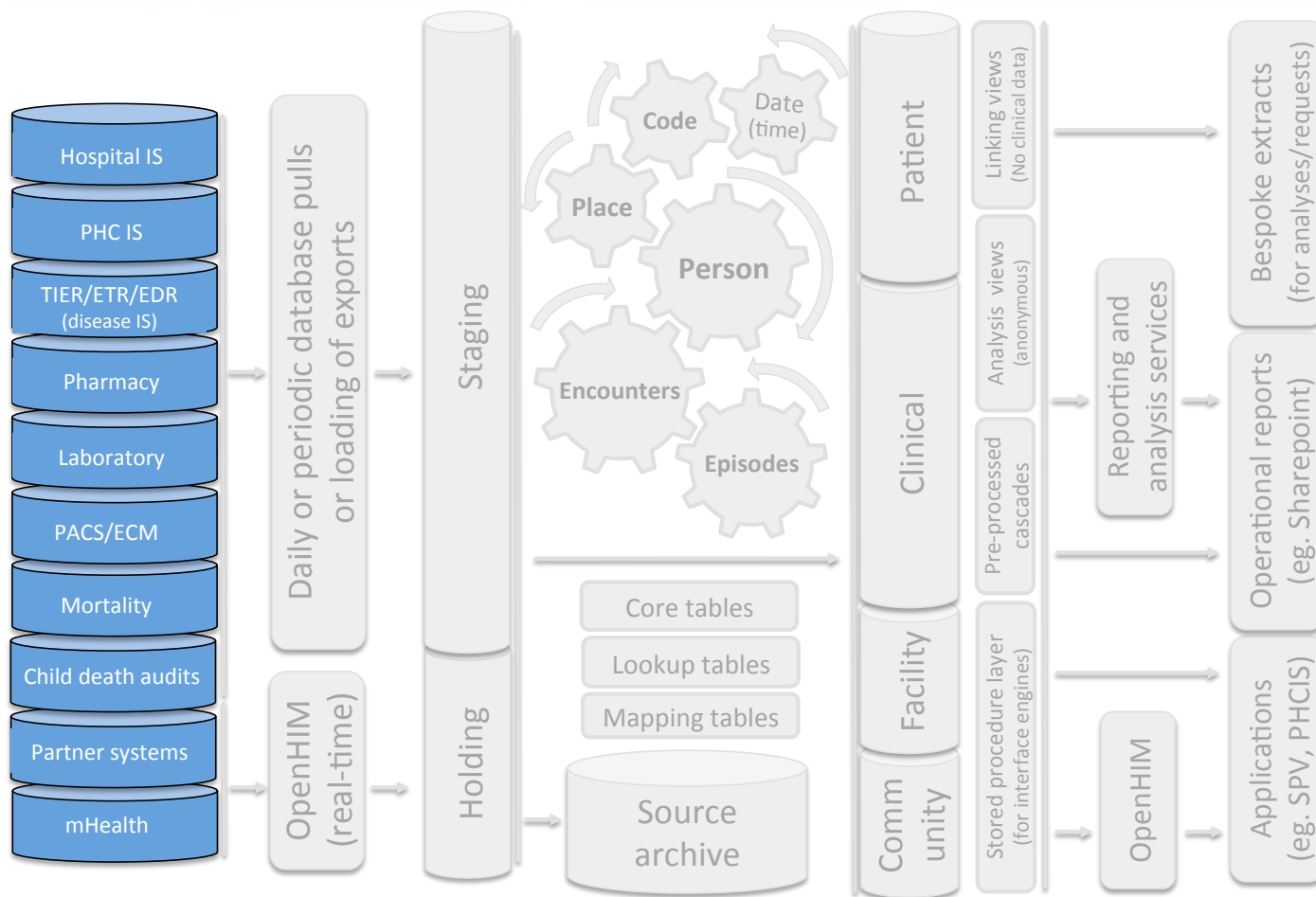
Proposed roles of a PHDC (Healthcare 2030)



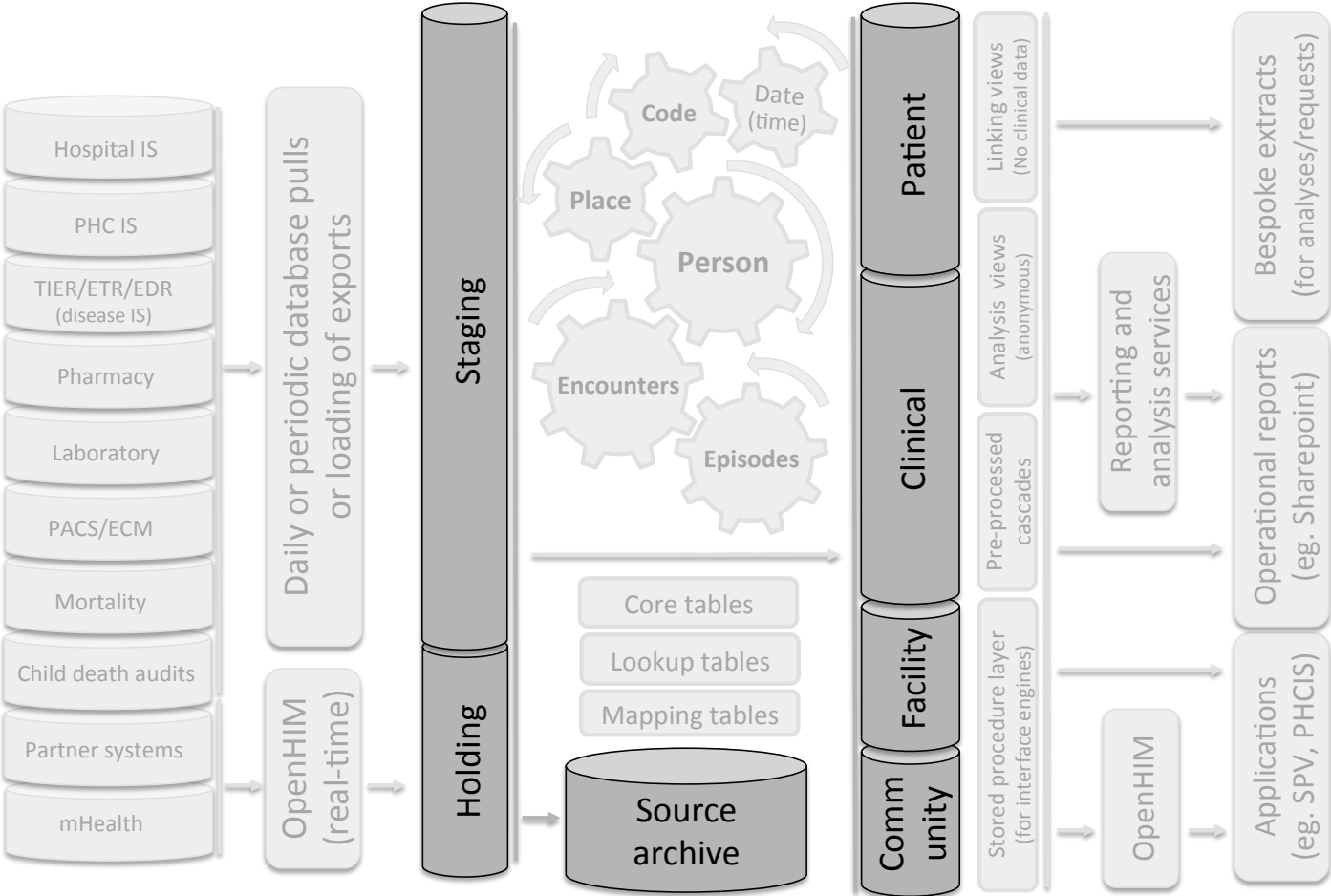
High level architecture



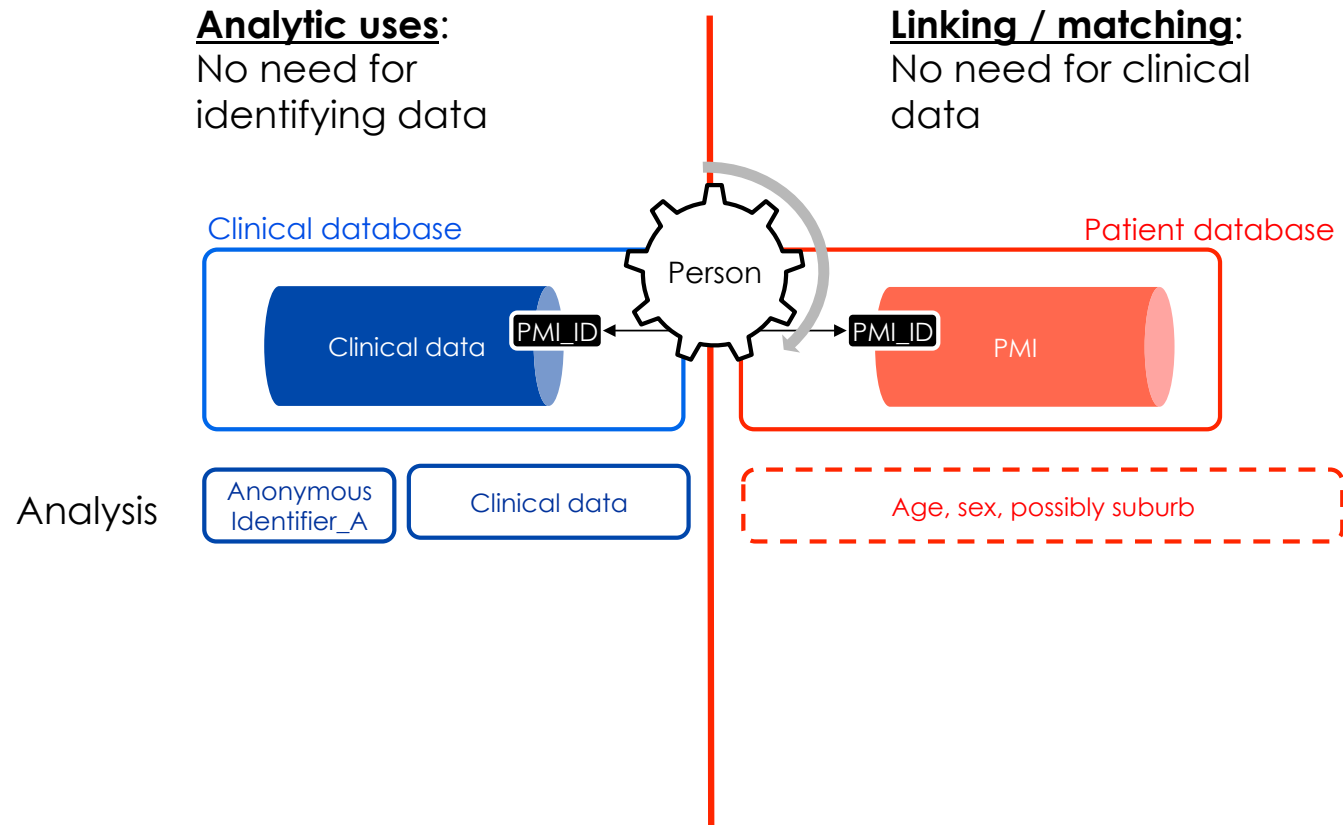
Source systems



PHDC core databases



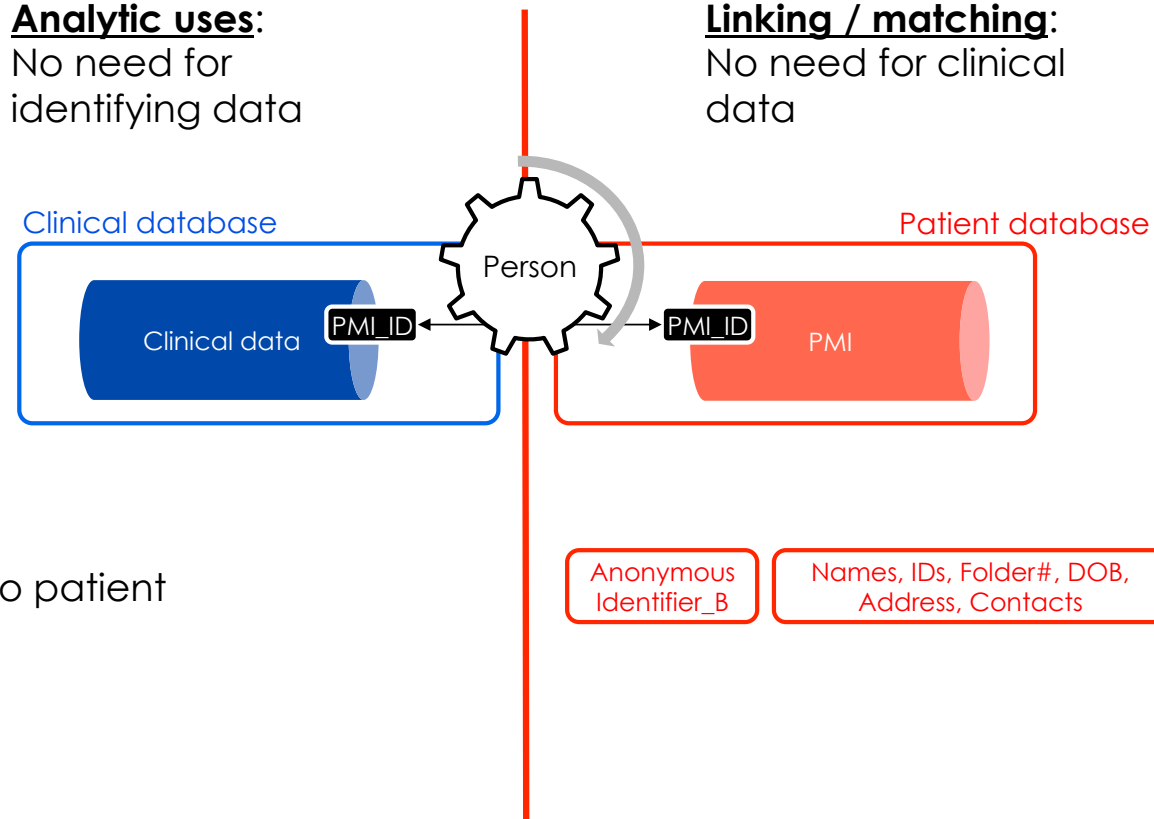
Separation of identifying and clinical data



Separation of identifying and clinical data

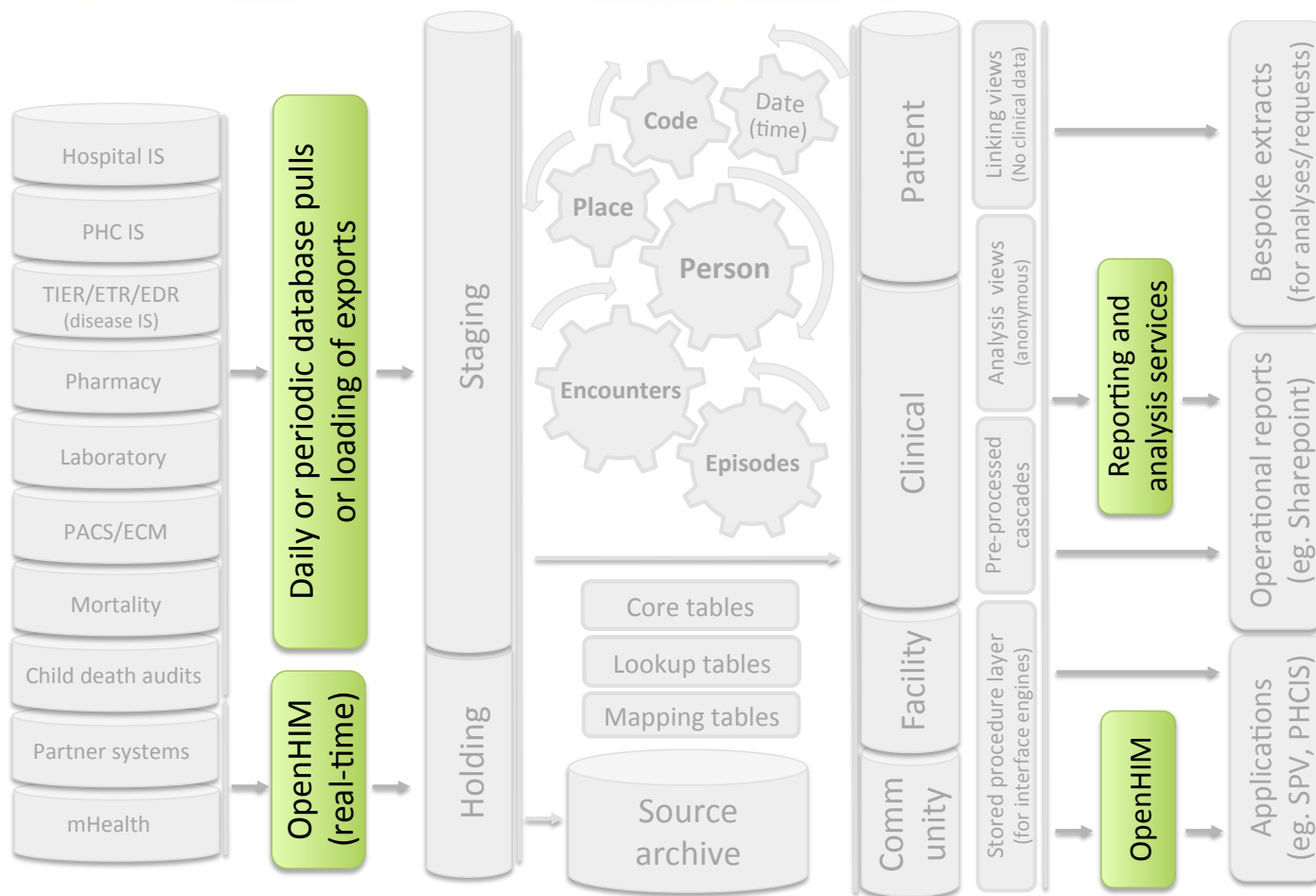
Analytic uses:
No need for
identifying data

Linking / matching:
No need for clinical
data

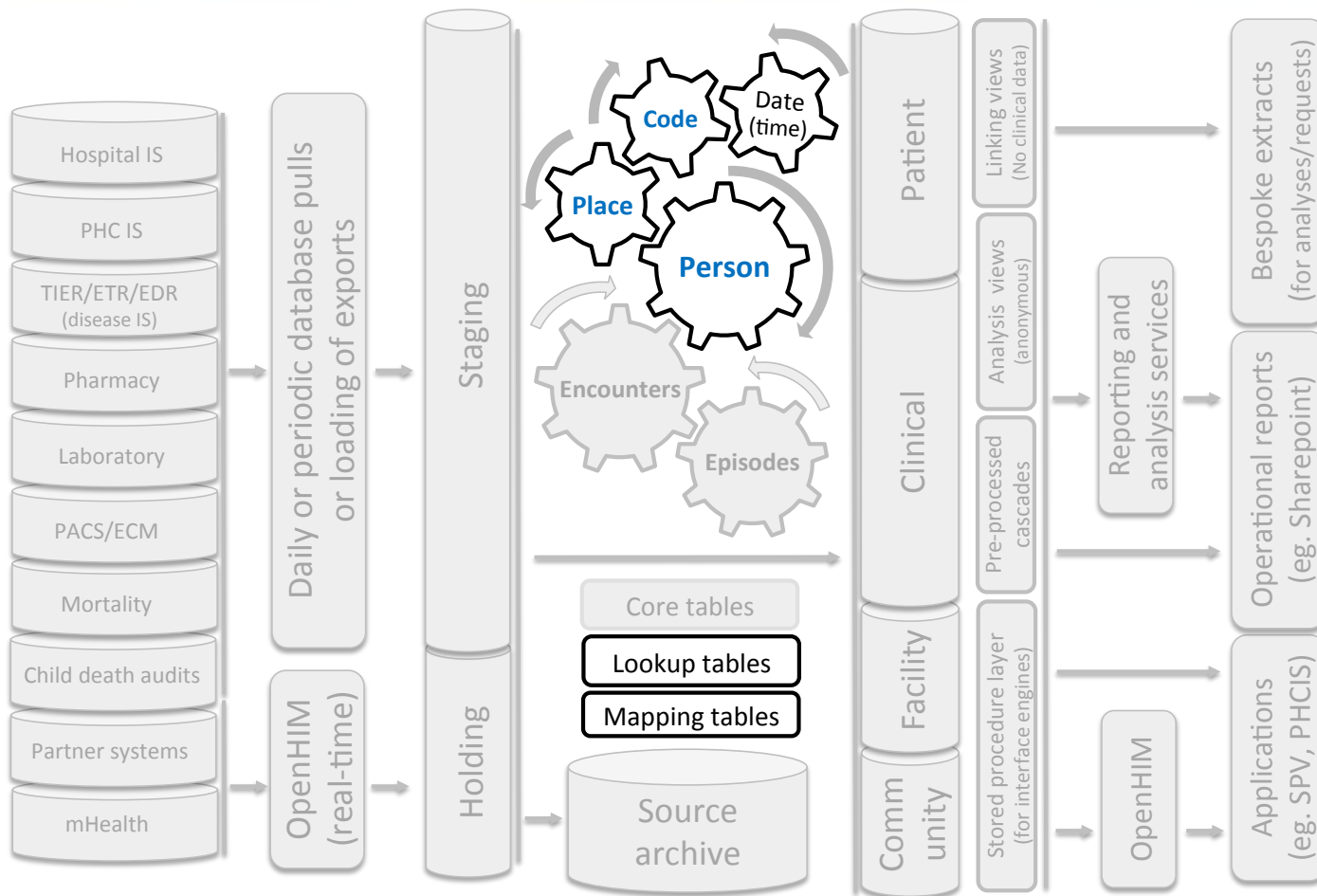


Linkage to patient

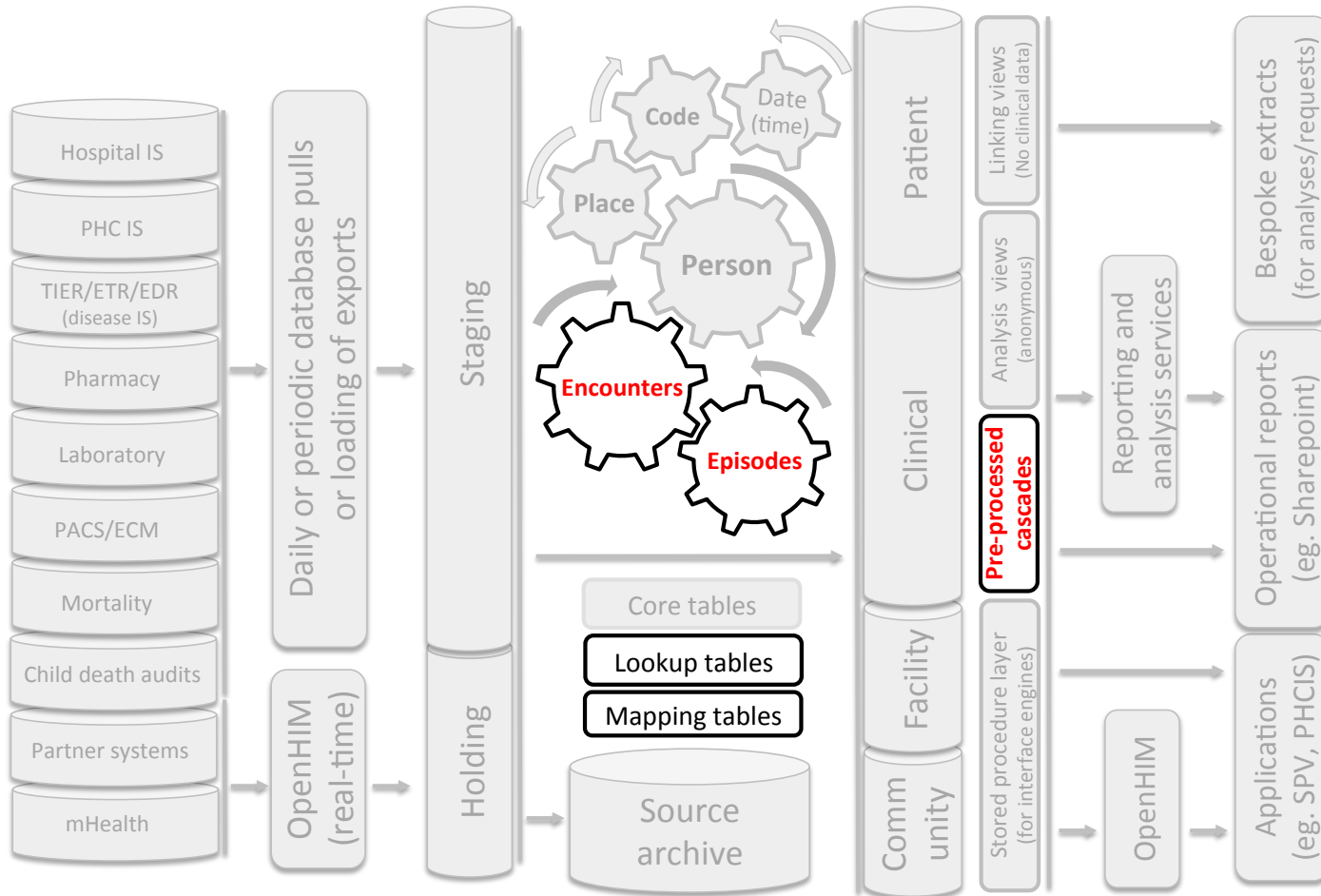
Tools for data take-on and access



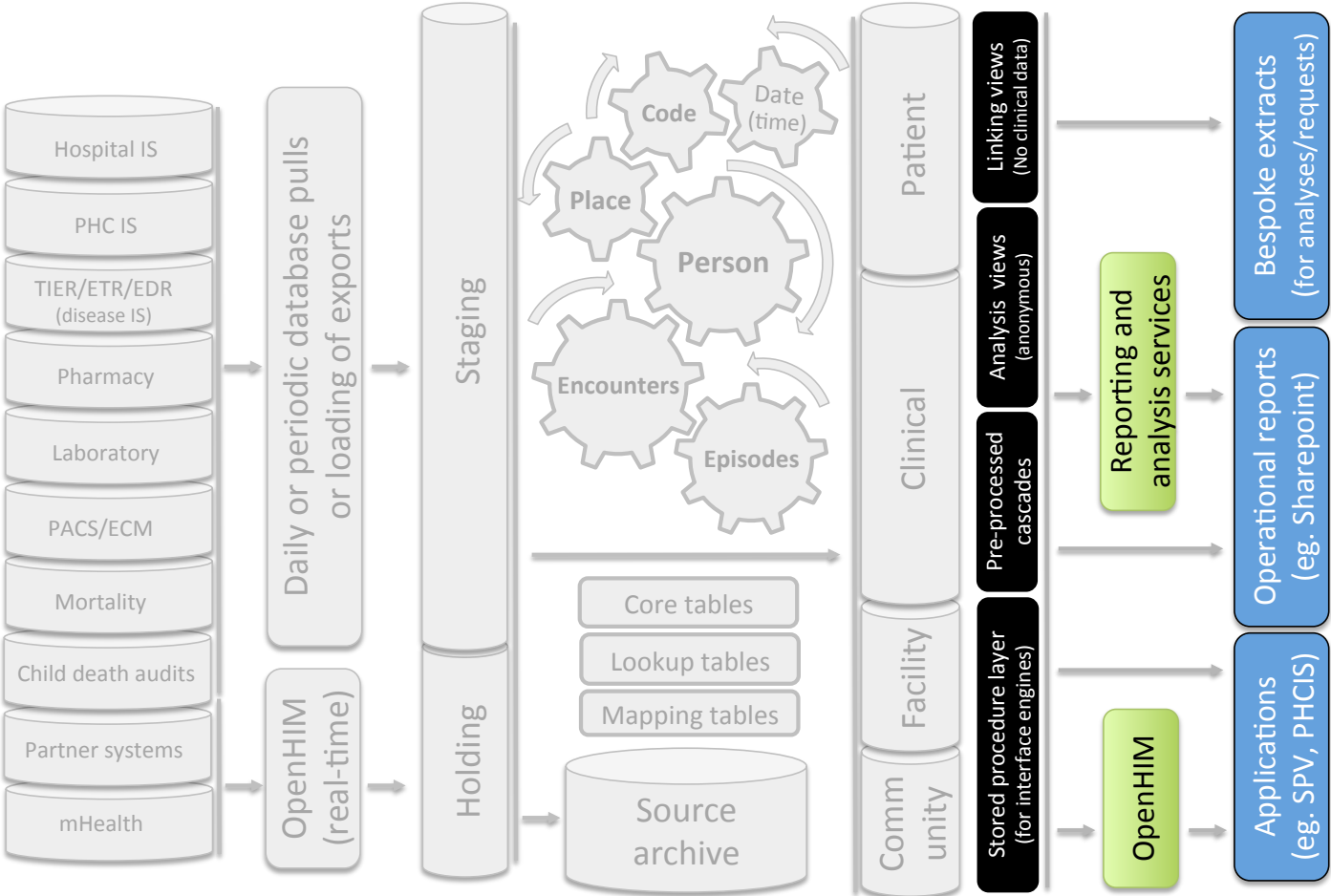
Data curation



Data beneficiation / enrichment



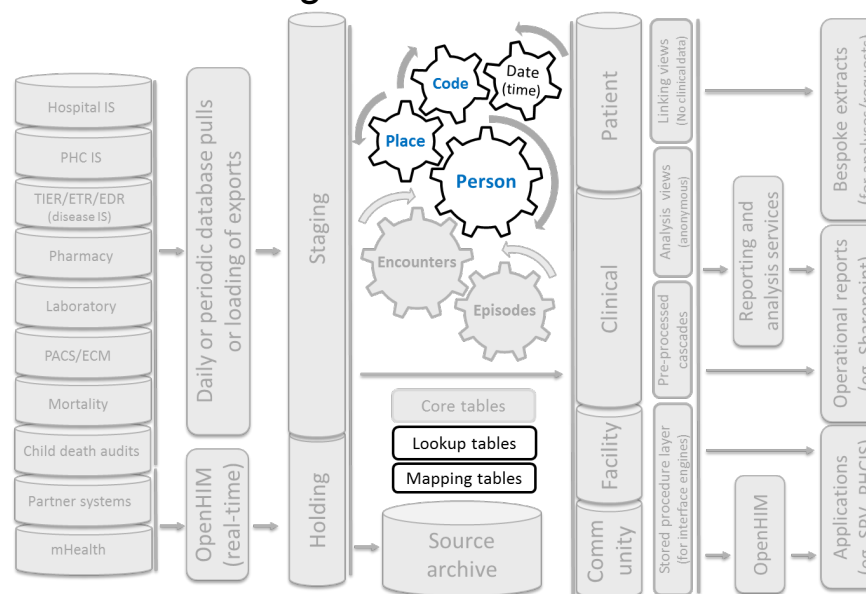
Data access and reporting



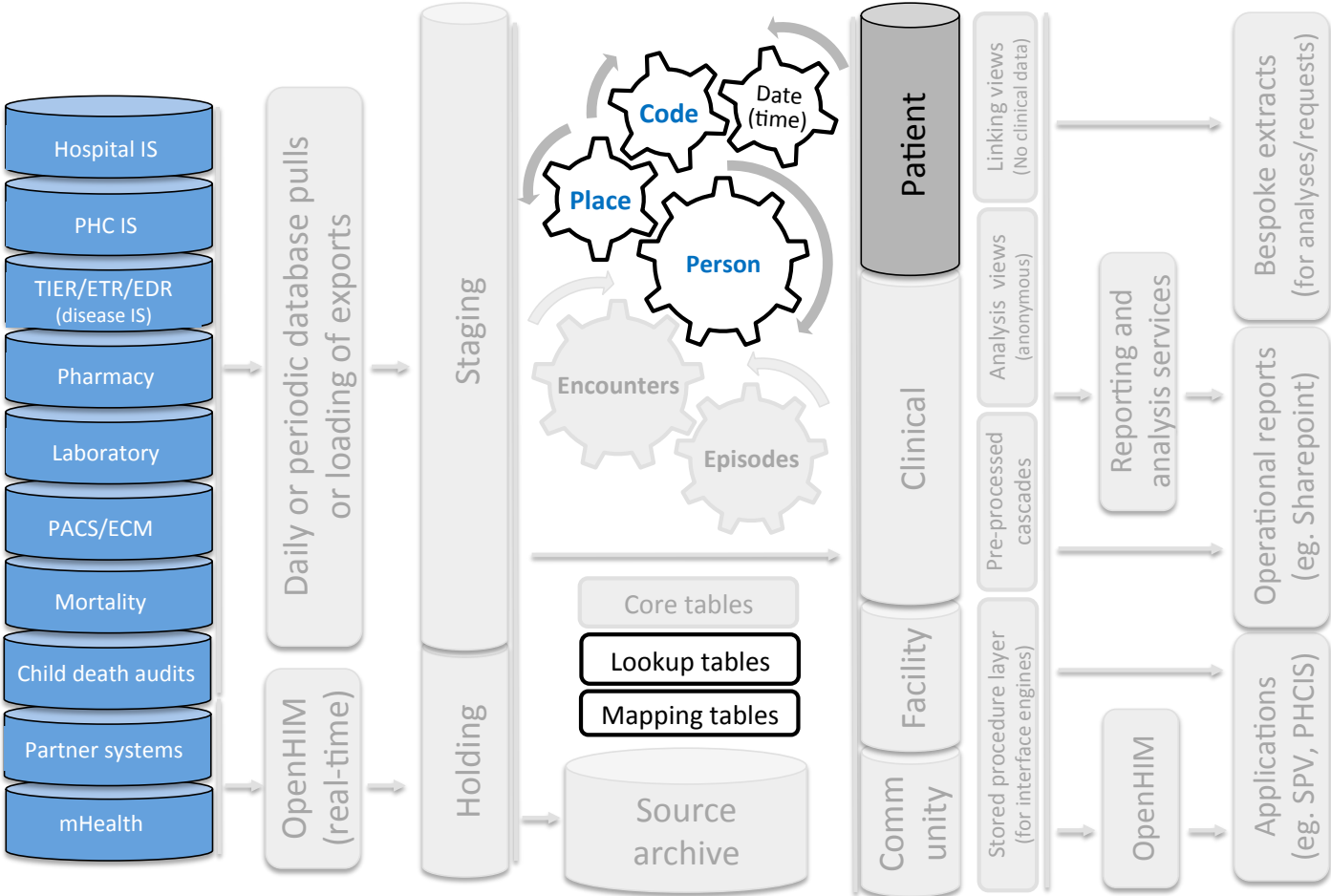
Outline

- ❑ Introduction, background, architecture
- ❑ **Data curation**
- ❑ Data beneficiation and enrichment: encounters, episodes and cascades
- ❑ Reporting
- ❑ Data governance, sharing and national alignment

- PMI and patient details
- Vital data
- Facilities
- Laboratory results
- Pharmacy records
- Diseases information systems
- Other sources



PMI, patient details, addresses, vital data



Daily PMI load (2017-04-20)

PMI		
Subsection	Query	Result
TIME	Time started	2017-04-20 00:05:54
	Time finished	2017-04-20 00:22:40
DATES	Earliest date extracted	2017-04-19 00:05:00
	Latest date extracted	2017-04-19 23:54:00
ADDITIONS	Number of patients added to the PMI from CLINICOM	1 948
	Number of New PMI records from PHCIS	1 487
MERGES	Permanent number merges	91
	Temporary number merges	230
	Permanent merges already linked	27
REVISIONS	Number of revised records from CLINICOM	1 251

17 minutes to run

Refers to all data added to Clinicom on 2017-04-19

New folder numbers (PHCIS only added to history)

Revisions to patient details

PMI - New patients added by source system			
Source system	Total patients	Probable duplicates	Total excluding duplicates
Prehmis	376	78 (21%)	298
PHCIS	749	95 (13%)	654
Clinicom	574	65 (11%)	509
Clinicom temporary PMI	244	15 (6%)	229
Eka	3	0 (0%)	3
PMI details not available	2	0 (0%)	2
	1948	253 (13%)	1695

Total number of folders created by source system & duplicate %

PMI | Matching data from Clinical sources

- We match patient data from the Clinical sources to the PMI
- Probable matches:
 - folder number + at least 1 other patient identifier
 - Or 3 or more exact matches on other patient identifiers
- Possible matches:
 - 3 highly similar matches on the other patient identifiers
- We store every identifier received from the source systems with their match

Example of daily patient numbers from NHLS (2017-04-19)

		Top 10 non-pmi facilities	
Number of distinct patients	13 278	43 : 2 Military Hospital	28
Number of patients mapped using pmi_links history (previously mapped)	12 192 (91.82 %)	9000001 : UCT Private Academic Hospital	23
Number of patients with exact matches	735 (5.54 %)	635 : Pollsmoor Correctional Centre	19
Number of patients with highly probable matches	41 (0.31 %)	829 : Vredenburg Hospital	18
Number of patients with possible matches	71 (0.53 %)	9000007 : Environmental Health	9
Number of unmapped patients	239 (1.80 %)	282 : Goodwood Correctional Centre	6
Number of unmapped patients by environmental samples	16 (6.69 %)	9000005 : SA Bone Marrow Registry	4
Number of unmapped patients by PMI facility	83 (34.73 %)	1342 : George Sickbay (SANDF) Clinic	3
Number of unmapped patients by non-PMI facility	149 (62.34 %)	333 : Helderstroom Correctional Centre	3
		2446 : Malmesbury CDC	3

PMI | Patient information



Names, Date of birth, Sex, National ID, Passport #



Addresses (Geocoding in process thanks to Robin & Global Fund)



Contacts



Folder locations



Meta data (e.g. # folders registered, last activity dates)



Family relationships

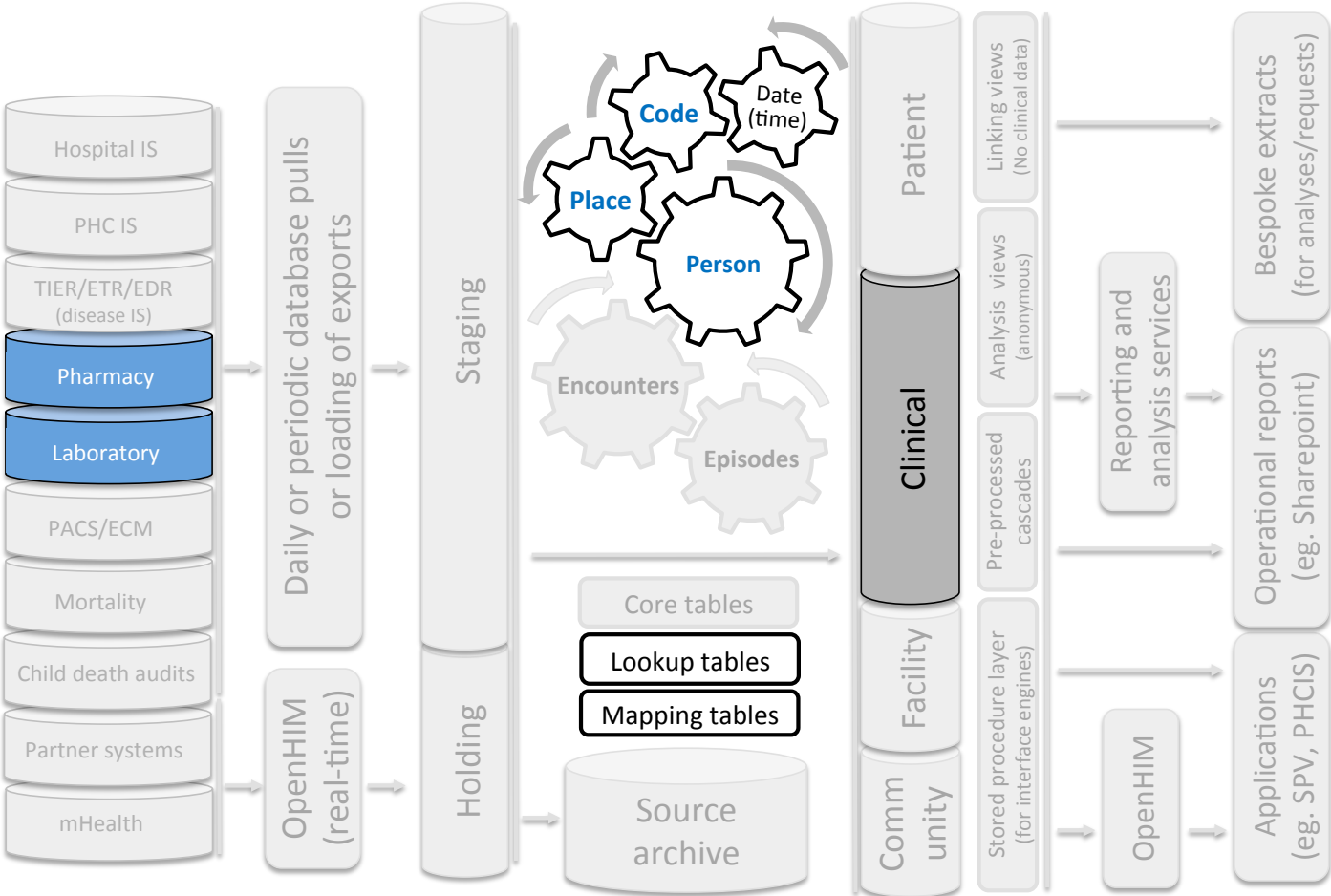


Births



Mortality

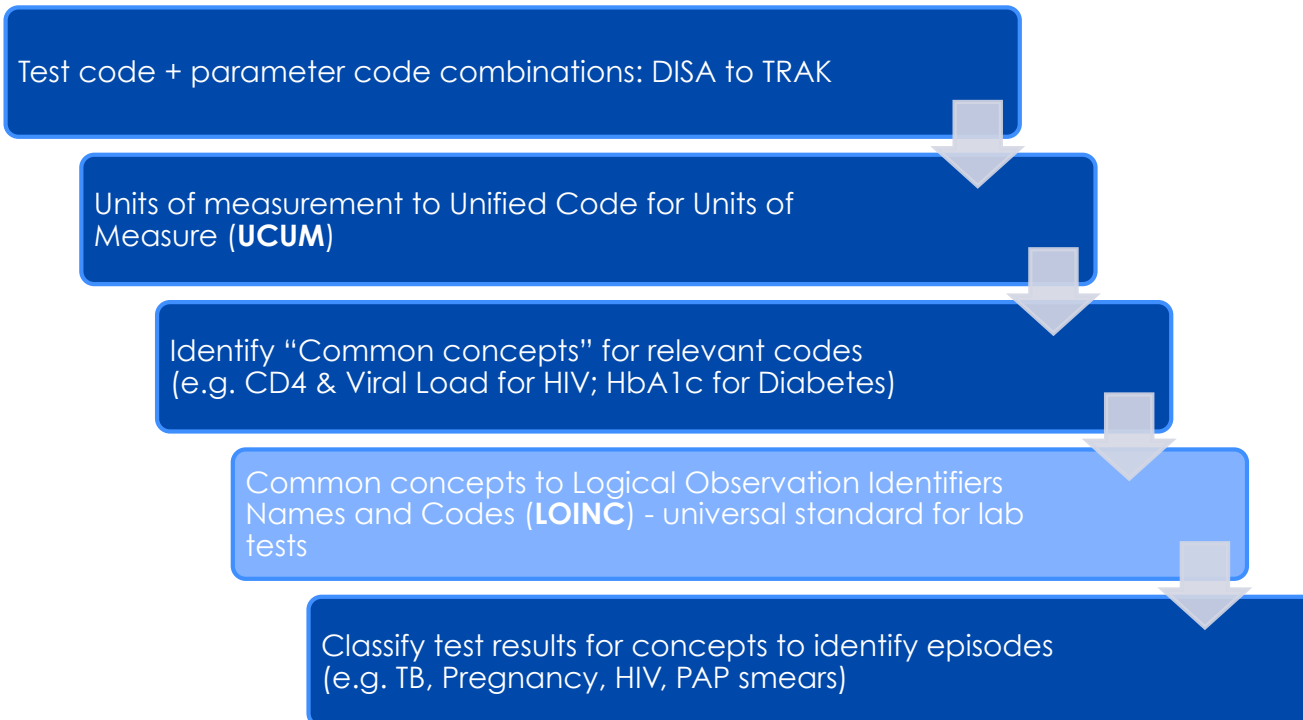
Laboratory and pharmacy data



NHLS Test Code Mapping

- ❑ TRAK and DISA: 17,037 test and parameter code combinations

What gets mapped?

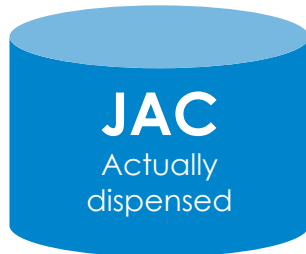


Common Concept Example: TB Microscopy

Test Code	Test Description	Parameter Code	Parameter Description	Common Group	Common Concept	Common Method
TBA	TB MICROSCOPY	TBACO	RESULT	TB	Microscopy Result	Auramine
TBA	TB MICROSCOPY	TBAC1	NOT AVAILABLE	TB	Microscopy Result	Auramine
TBCUL	TUBERCULOSIS INVESTIGATION	TBAUR	AURAMINE	TB	Microscopy Result	Auramine
M176	AURAMINE (WITH TB CULTURE)	M0025	AFB (CONCENTRATED)	TB	Microscopy Result	Auramine
M176	AURAMINE (WITH TB CULTURE)	M0024	AFB (RAPID UNCONCENTRATED)	TB	Microscopy Result	Auramine
M005	AURAMINE O STAIN	M0024	AFB (RAPID UNCONCENTRATED)	TB	Microscopy Result	Auramine
M005	AURAMINE O STAIN	M0025	AFB (CONCENTRATED)	TB	Microscopy Result	Auramine
TBZ	TB MICROSCOPY	TBZCO	ZN RESULT	TB	Microscopy Result	Ziehl-Neelsen
TBCUL	TUBERCULOSIS INVESTIGATION	TBZN	ZIEHL - NEELSEN	TB	Microscopy Result	Ziehl-Neelsen
M195	AFB (+VE RE-DECONTAMINATE)	M0201	ZN RESULT (POSITIVE CULTURE)	TB	Microscopy Result	Ziehl-Neelsen
M185	AFB (POSITIVE MGIT CULTURE)	M0201	ZN RESULT (POSITIVE CULTURE)	TB	Microscopy Result	Ziehl-Neelsen
M185	AFB (POSITIVE MGIT CULTURE)	M0252	ZN RE-INCUBATED	TB	Microscopy Result	Ziehl-Neelsen
B160	NICD: AFB (POSITIVE CULTURE)	M0201	ZN RESULT (POSITIVE CULTURE)	TB	Microscopy Result	Ziehl-Neelsen
B161	NICD: DST (POSITIVE CULTURE)	B0065	ZN RESULT	TB	Microscopy Result	Ziehl-Neelsen
M181	TB CULTURE SOLID MEDIUM	M0201	ZN RESULT (POSITIVE CULTURE)	TB	Microscopy Result	Ziehl-Neelsen
M181	TB CULTURE SOLID MEDIUM	M0252	ZN RE-INCUBATED	TB	Microscopy Result	Ziehl-Neelsen
M010	ZIEHL-NEELSEN STAIN	M0024	AFB (RAPID UNCONCENTRATED)	TB	Microscopy Result	Ziehl-Neelsen
M010	ZIEHL-NEELSEN STAIN	M0025	AFB (CONCENTRATED)	TB	Microscopy Result	Ziehl-Neelsen

Pharmacy

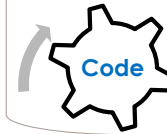
Current datasources included in our pharmacy table



Prospective datasources to be added

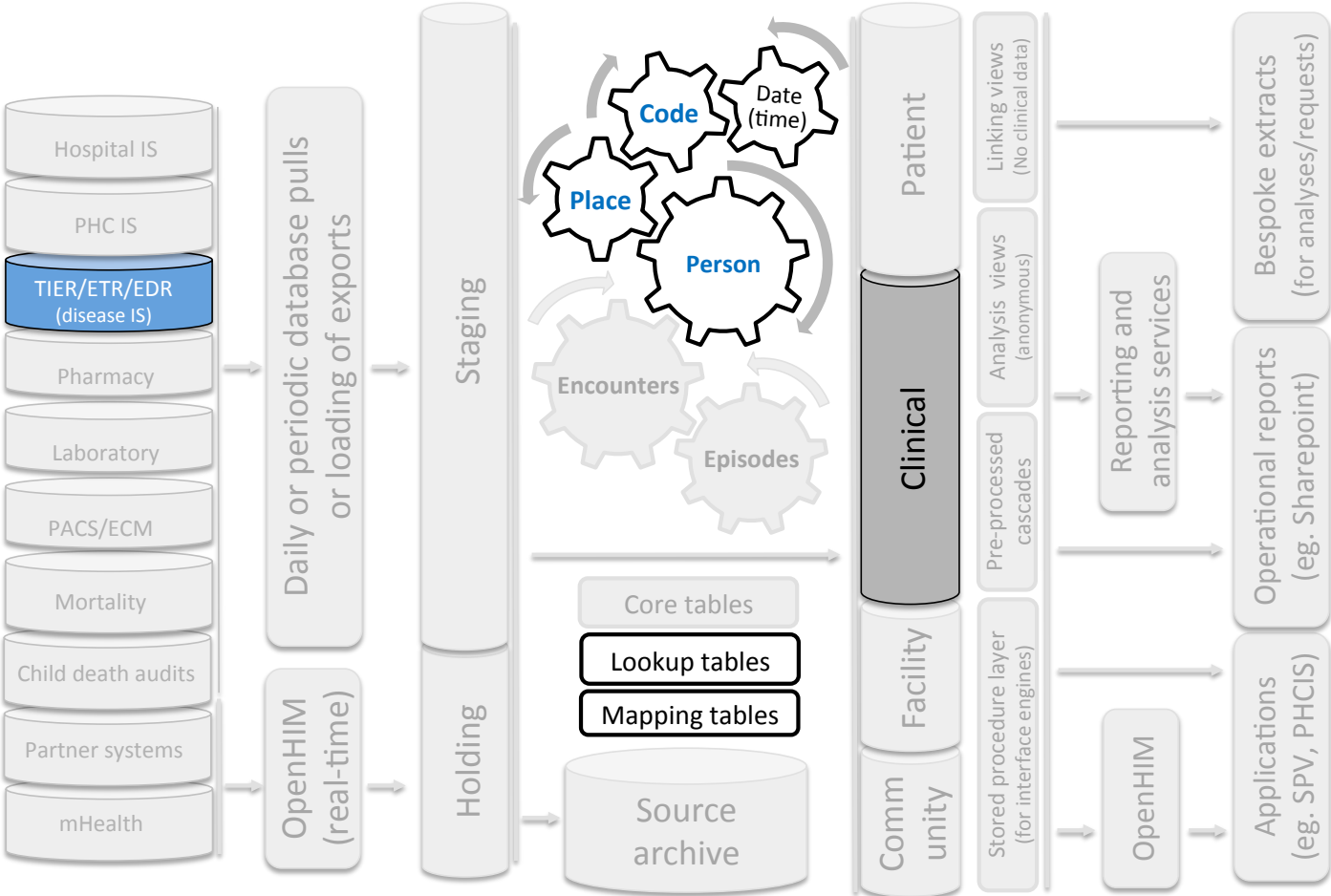


All codes are mapped to a common system



WHO ATC
(Anatomical Therapeutic Chemical
(ATC) classification system)

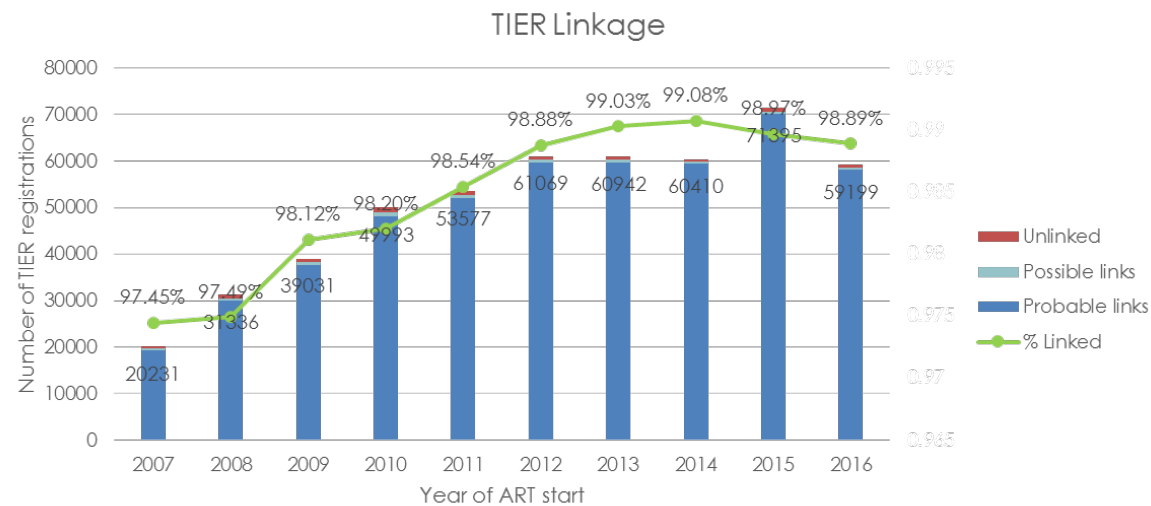
Disease information systems



Disease Information Systems

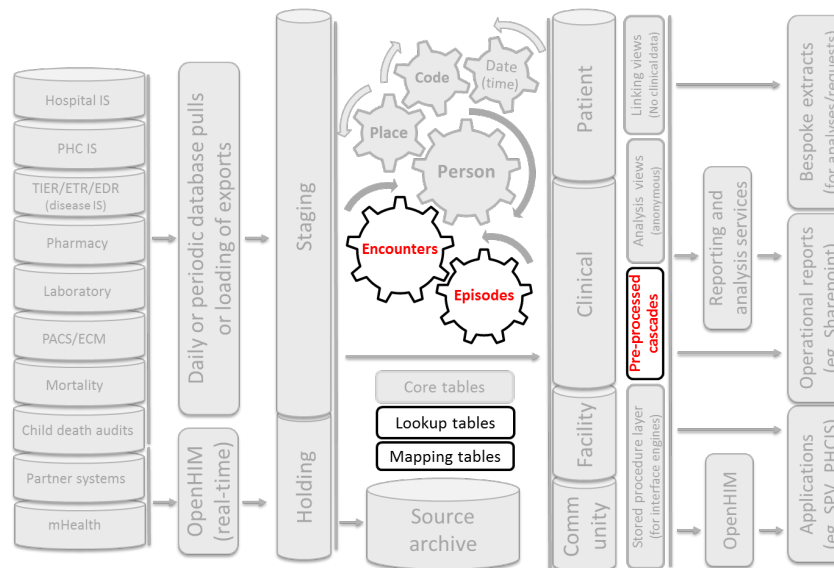


- DIS data is received quarterly
- When received it is incorporated into existing data structures to supplement data received daily
- As the coverage of PHCIS/EKAPA increases we will start to receive the DIS data more real time



Outline

- ❑ Introduction, background, architecture
- ❑ Data curation
- ❑ **Data beneficiation and enrichment: encounters, episodes and cascades**
- ❑ Reporting
- ❑ Data governance, sharing and national alignment



Encounters

- Facility visit ≈ Encounter



- Encounter types & Sources:



Admissions: One line per admission with admission & discharge date [CLINICOM]



OP Hospital Visits: One line per OP clinic, date & time [CLINICOM]



PHC Visits: One line per patient, facility & date (time not reliable) [PHCIS/EKAPA/PREHMIS*]



DIS Visits: One line per patient, facility & date. Inferred visits from DIS treatment dates. [TIER/ETR]



Dispensing: One line per patient, facility & date. Inferred visits from drug dispensing. [JAC]



Lab visits: One line per patient, facility & date. Inferred visits from taken labs. [NHLS]



Community: Self enrolment dates [MomConnect]/Community registrations [Catch&Match]

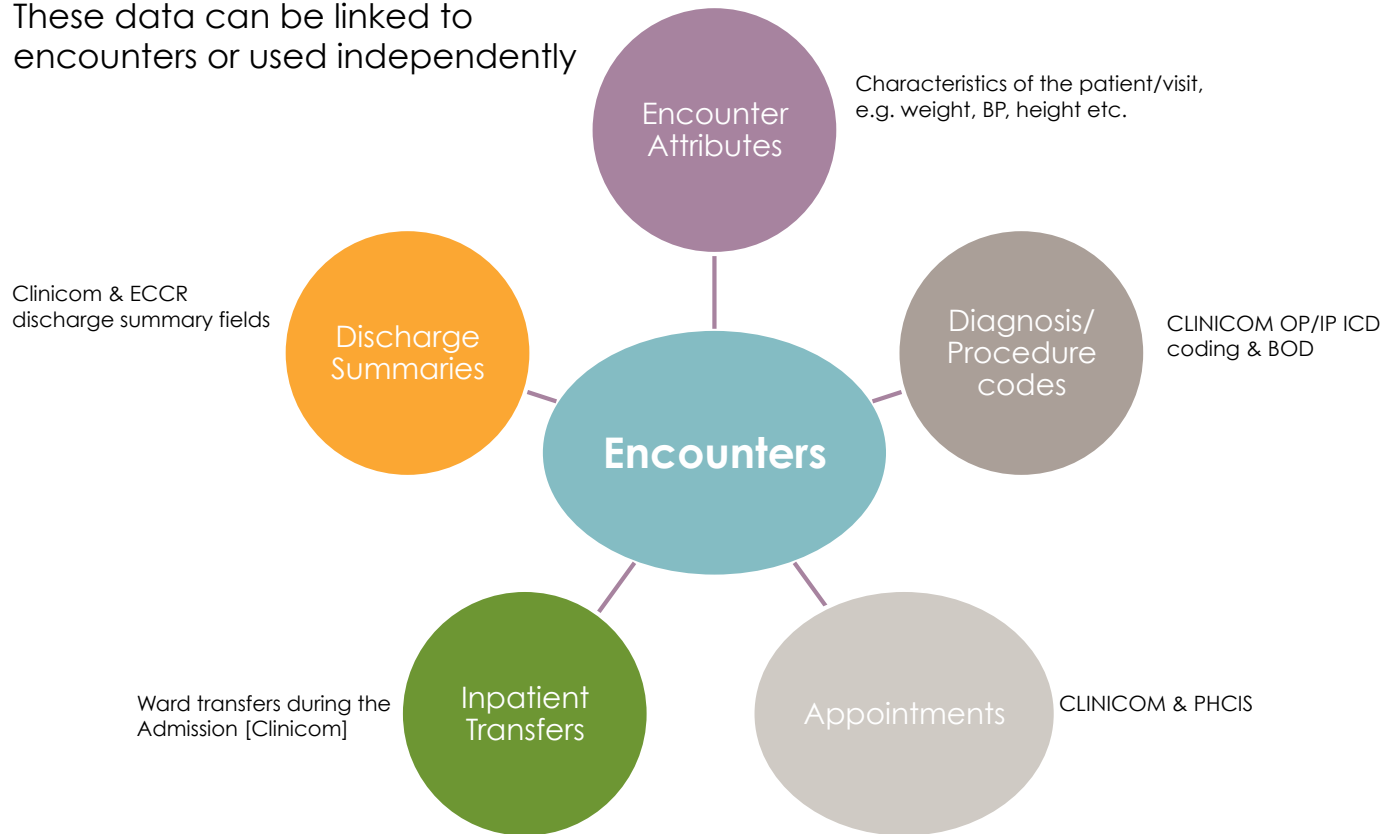
Daily encounters (2017-04-20)

ADMISSIONS	Number of open admissions that were closed	0
	Number of new admissions added	1 969
OP VISITS	Number of new OP visits added	12 651
	Number of OP visits updated	0
PHC VISITS	Total number of new PHC visits	30 139
PHCIS VISITS	Transaction date loaded from	2017-04-20 00:15:11
	Transaction date loaded to	2017-04-20 00:15:14
	Earliest visit date	2017-03-17 00:00:00
	Latest visit date	2017-04-19 00:00:00
	Number of new PHCIS visits added	18 314
	Number of visits updated to include PHCIS visits	2 346
PHCIS FIRST VISITS	Transaction date loaded from	2017-04-20 00:15:18
	Transaction date loaded to	2017-04-20 00:15:18
	Earliest visit date	2017-04-19 00:00:00
	Latest visit date	2017-04-19 00:00:00
	Number of new PHCIS First visits added	2 107
	Number of visits updated to include PHCIS FIRST visits	348
EKAPA VISITS	Transaction date loaded from	2017-04-20 00:15:44
	Transaction date loaded to	2017-04-20 00:15:44
	Earliest visit date	2016-04-05 00:00:00
	Latest visit date	2017-04-19 00:00:00
	Number of new EKAPA visits added	861
	Number of visits updated to include EKAPA visits	392
JAC DISPENSING	Number of new JAC dispensing encounters added	7 783
	Number of visits updated to include JAC dispensing	3 853
LABS	Number of new lab encounters added	1 074
	Number of visits updated to include labs	2 248

Encounter Type	Count
Admissions	1969
OP Visits	12651
PHC Visits (Total)	30139
PHCIS Visits	18314 (60.8%)
PHCIS First Visits	2107 (7.0%)
Ekapa Visits	861 (2.9%)
JAC Dispensing	7783 (25.8%)
Labs	1074 (3.6%)

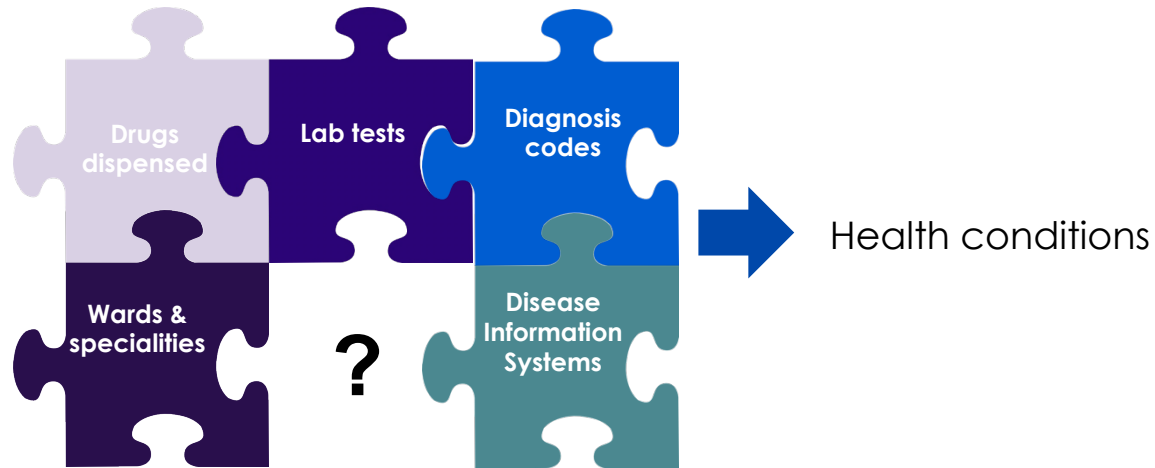
Information we link to encounters

These data can be linked to encounters or used independently



Episodes

- Evidence from clinical data enables us to infer health conditions



- Episodes can be acute/chronic
- We combine evidence from clinical data into a single record per patient per episode
- Includes: First & Last evidence dates, treatment start dates and outcome dates, as well as facilities, datasources & a list of evidence

Datasources used to ascertain pregnancy

SOURCE	EVIDENCE (bold=high confidence)
Birth records (CLINICOM, PHCIS and PPIP)	<ul style="list-style-type: none"> • Birth outcomes (e.g. live birth, still birth etc)
Laboratory (NHLS)	<ul style="list-style-type: none"> • Rhesus Testing, Pap Smear with pregnancy parameter • Supporting evidence: BETA HCG, Syphilis
Pharmacy (JAC)	<ul style="list-style-type: none"> • Terminations by mifepristone • Supporting evidence: Iron & Folate, Metaclopramide
Admissions & OP visits (CLINICOM)	<ul style="list-style-type: none"> • ICD10 codes • Supporting evidence: Ward or OP clinic type, Speciality codes, internal codes describing admission/discharge methods
Community (MOM CONNECT, CATCH&MATCH*)	<ul style="list-style-type: none"> • Mom connect registrations
PHC Visits (PHCIS, EKAPA, PREHMIS*, inferred visits from JAC/NHLS)	<ul style="list-style-type: none"> • MOU Visits (weak-moderate)

* To be added soon

Episode evidence categories

Outcome	<ul style="list-style-type: none">• Evidence of an outcome for the episode• High confidence score
High confidence	<ul style="list-style-type: none">• Strong enough evidence to start a pregnancy episode• High confidence score
Weak-Moderate evidence	<ul style="list-style-type: none">• Able to start an episode, however more evidence is required to improve confidence in the episode• $\geq 5 \rightarrow$ High confidence score.• $< 5 \rightarrow$ Low confidence score.
Supporting only	<ul style="list-style-type: none">• Can only be appended to existing episode• It is not strong enough to start an episode• Improves confidence in existing episodes

Inferring pregnancy episodes

Example evidence list

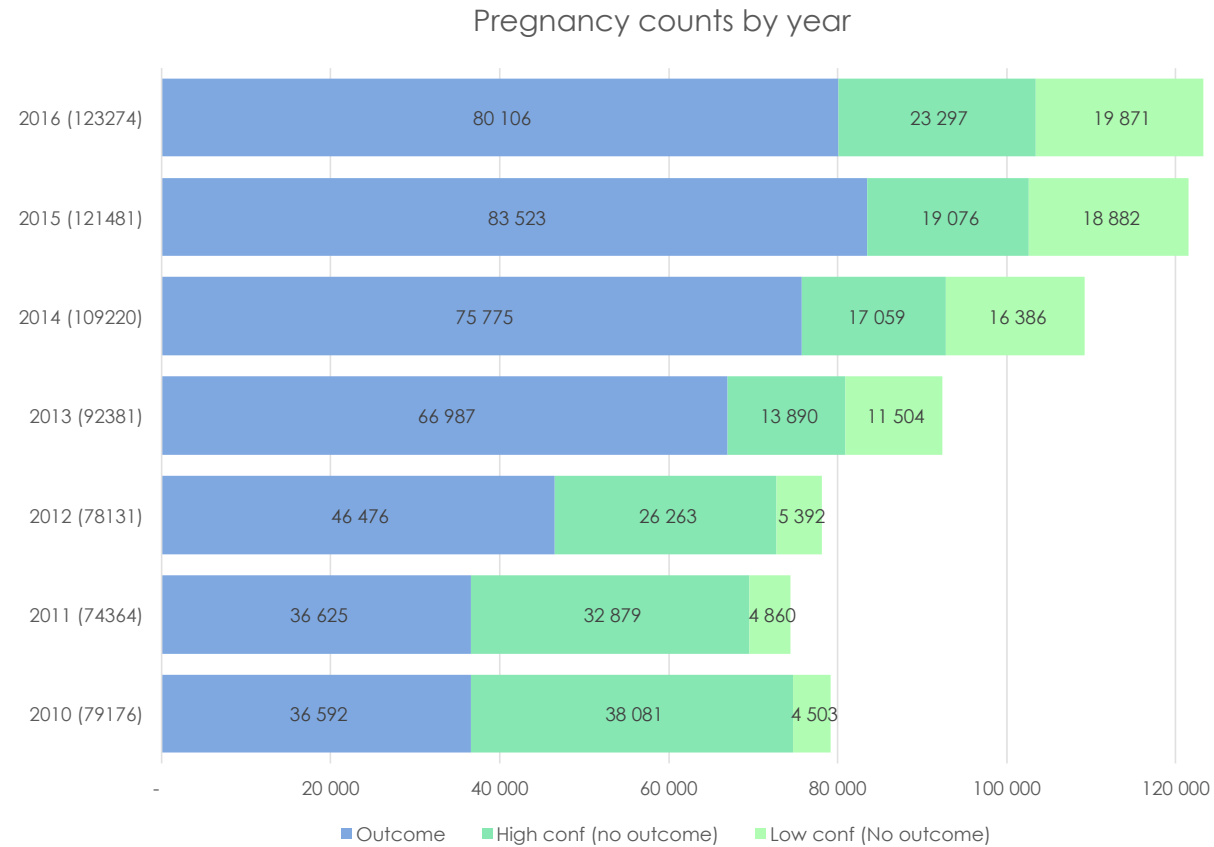
Evidence	Source	Category	Date
Rhesus test performed	Labs	High confidence	2015-01-23
Iron & Folate dispensed	Drugs	Supporting	2015-04-03
Has live birth record	Birth register	Outcome	2015-06-16
Has diagnosis code indicating live birth	Diagnosis codes	Weak-Moderate	2015-06-16



Rolled up into a single record per pregnancy
(using a sliding date range of at most 270 days between the first and last contact date)

Patient	Pregnancy number	Start Date	End Date	Last Contact date	Evidence list	End date evidence list	Facility	Confidence
xxx	1	2015-01-23	2015-06-16	2015-06-16	Birth Record, Diagnosis code, Rhesus Test, Iron & Folate	Birth Record, Diagnosis code	MMH	0.95 → High confidence

Estimated pregnancies 2010 - 2016



The number of pregnancies with the last date of contact within the given year

Rolling Episode counts (as of 2017-01-01)

Episode Type	High confidence episodes [Total*]
HIV positive	347 550 [412 370]
HIV positive on ART	245 024
Diabetes Mellitus	246 387 [246 896]
Hypertension	407 738
Asthma or COPD	187 344
Epilepsy	30 471
Pregnancy	96 813 [114104]
TB positive	52 088 [58 646]
TB on treatment	33 443 *[39816 with extra rollback to 2016-06-01]

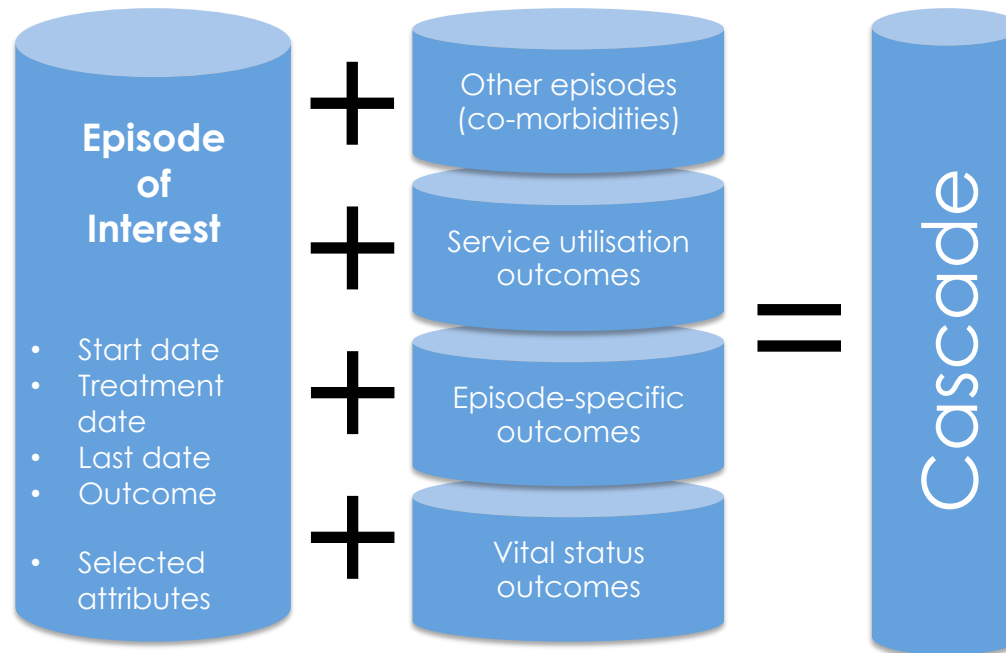
*All episodes where the high confidence count is different from the total count

Note that for chronic episodes, counts are calculated based on patients who have no death record and have had any encounter within 2 years before the report date. For HIV, the dates were shifted earlier by 6 months to account for the lag in disease information systems. For HIV, the number on treatment is based on those with an ART start date and any encounter with TIER/EKAPA in the last 18 months.

For Pregnancy & TB, counts are calculated based on patients who started their episode within 1 year before the report date. For TB, the dates were shifted earlier by 6 months to account for the lag in disease information systems. For TB, the number on treatment is based on those registered in EDR/ETR.

Cascades

- Comprehensive patient level view of episodes, including key dates, evidence, outcomes and co-morbidities relevant for patient management
- Provides us with a cohort in the absence of complete registers
- Optimised for operational and management purposes
- Easily aggregated by region or attributes

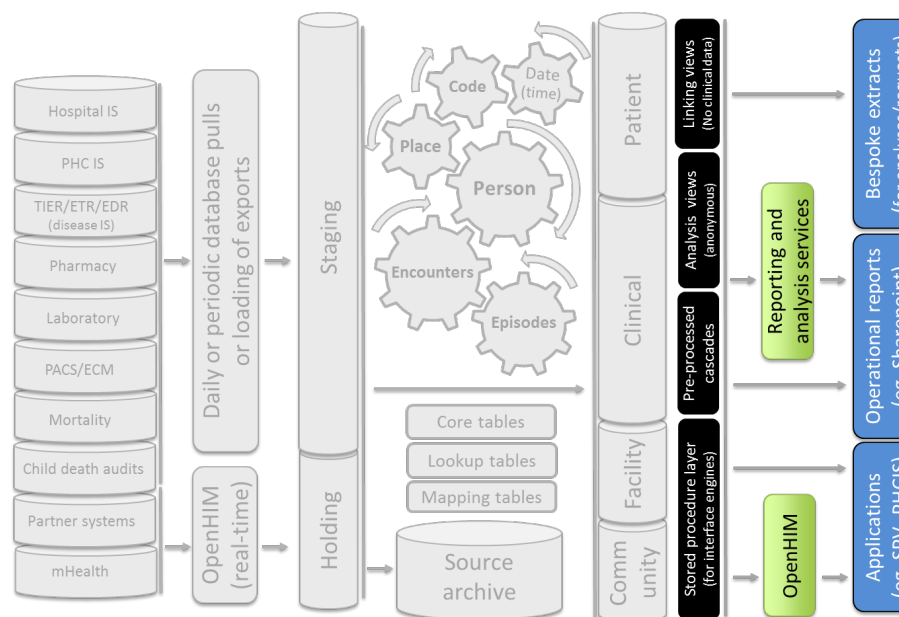


Cascades | Current & future

- Maternal and infant
- HIV
- TB
- Diabetes
- Cervical health (Pap smear & Colposcopy)
- Antimicrobial resistance / Hospital acquired infections
- (Notifiable conditions)

Outline

- ❑ Introduction, background, architecture
- ❑ Data curation
- ❑ Data beneficiation and enrichment: encounters, episodes and cascades
- ❑ **Reporting**
- ❑ Data governance, sharing and national alignment



Sample output – all pregnancies in 2015, n=106588, limited to high confidence

Health BI > Clinical BI > Data Centre > Reports > Confidential

Actions 2221 of 2221 Find Next

Row ID	Folder No	Age	Pregnancy Evidence	Valid	HIV status	Dead	Co-morbidities	Pregnancy First Evidence Date	EDD	Pregnancy Outcome Date	Infant Outcome	Infant HIV Status	Birth Weight (g)
106569	155xxxxxxxx	29	BIRTH OUTCOME	✓	UNKNOWN			2015/12/30	2016/03/27	2016/03/27	Live Birth	UNKNOWN	3210
106570	105xxxxxxxx	23	BIRTH OUTCOME	✓	UNKNOWN			2015/12/30	2015/12/30	2015/12/30	Live Birth	UNKNOWN	2540
106571	155xxxxxxxx	23	BIRTH OUTCOME	✓	UNKNOWN			2015/12/30	2015/12/30	2015/12/30	Live Birth	UNKNOWN	2990
106572	105xxxxxxxx	32	BIRTH OUTCOME	✓	ART during (2015/08/31)			2015/12/31	2016/01/01	2016/01/01	Live Birth	UNKNOWN	2800
106573	112xxxxxxxx	27	BIRTH OUTCOME	✓	UNKNOWN			2015/11/06	2016/01/23	2016/01/23	Live Birth	UNKNOWN	3740
106574	806xxxxxxxx	24	BIRTH OUTCOME	✓	UNKNOWN		Asthma or COPD	2015/12/15	2016/04/26	2016/04/26	Live Birth	UNKNOWN	3565
106575	200xxxxxxxx	29	BIRTH OUTCOME	✓	UNKNOWN			2015/12/15	2015/12/15	2015/12/15	Live Birth	UNKNOWN	2700
106576	573xxxxxxxx	35	BIRTH OUTCOME	✓	UNKNOWN			2015/12/15	2015/12/15	2015/12/15	Live Birth	UNKNOWN	2700
106577	119xxxxxxxx	22	BIRTH OUTCOME	✓	UNKNOWN			2015/12/15	2016/05/22	2016/05/22	Live Birth	UNKNOWN	3520
106578	110xxxxxxxx	24	BIRTH OUTCOME	✓	UNKNOWN			2015/12/15	2016/07/09	2016/07/09	Live Birth	UNKNOWN	3760
106579	838xxxxxxxx	30	BIRTH OUTCOME	✓	UNKNOWN			2015/12/15	2016/05/04	2016/05/04	Live Birth	UNKNOWN	2960
106580	149xxxxxxxx	20	HIGH CONFIDENCE	✓	UNKNOWN			2015/12/15	2016/01/12				
106581	441xxxxxxxx	18	BIRTH OUTCOME	✓	UNKNOWN			2015/12/15	2015/12/15	2015/12/15	Live Birth	UNKNOWN	2475
106582	150xxxxxxxx	19	BIRTH OUTCOME	✓	UNKNOWN		Hypertension	2015/12/15	2016/04/14	2016/04/14	Live Birth	UNKNOWN	3000
106583	842xxxxxxxx	17	BIRTH OUTCOME	✓	UNKNOWN			2015/12/16	2015/12/25	2015/12/25		UNKNOWN	3010
106584	132xxxxxxxx	30	BIRTH OUTCOME	✓	UNKNOWN			2015/12/16	2016/01/21	2016/01/21	Live Birth	UNKNOWN	2620
106585	785xxxxxxxx	23	BIRTH OUTCOME	✓	NEGATIVE (2015/12/16)			2015/12/16	2015/12/16	2015/12/16	Live Birth	UNKNOWN	1080
106586	155xxxxxxxx	18	BIRTH OUTCOME	✓	UNKNOWN			2015/12/16	2015/12/16	2015/12/16	Live Birth	UNKNOWN	3100
106587	155xxxxxxxx	29	BIRTH OUTCOME	✓	UNKNOWN			2015/12/17	2016/03/22	2016/03/22	Live Birth, Live Birth	UNKNOWN	4190, 4190
106588	683xxxxxxxx	30	BIRTH OUTCOME	✓	UNKNOWN			2015/12/17	2016/04/29	2016/04/29	Live Birth	NEGATIVE(2016-07-12)	2594

Sample output – maternal deaths in 2015, n=58

- overlap between death and pregnancy, **not** the institutional reporting
- only some will overlap with official institutional reporting

Health BI > Clinical BI > Data Centre > Reports > Confidential

Actions 2 of 2 Find Next

Western Cape Government
Western Cape Provincial Health Data Centre

Maternal Cascade Report
for pregnancies between 2015-01-01 and 2015-12-31
generated at 2017-04-05 08:57:16

Row ID	Folder No	Age	Pregnancy Evidence	Valid	HIV status	Dead	Co-morbidities	Pregnancy First Evidence Date	EDD	Pregnancy Outcome Date	Infant Outcome	Infant HIV Status	Birth Weight (g)
49	232xxxxxxx	37	BIRTH OUTCOME	✓	ART during (2015/12/11)	✓		2015/12/11	2016/02/26	2016/02/26	Live Birth	NEGATIVE(2016-02-26)	3145
50	569xxxxxxx	36	BIRTH OUTCOME	✓	UNKNOWN	✓		2015/10/17	2015/10/17	2015/10/17	Live Birth	UNKNOWN	1300
51	888xxxxxxx	37	BIRTH OUTCOME	✓	UNKNOWN	✓	Diabetes Mellitus, Mental Health	2015/10/20	2015/12/27	2015/12/27	Live Birth	UNKNOWN	3590
52	133xxxxxxx	24	BIRTH OUTCOME	✓	UNKNOWN	✓		2015/02/11	2015/09/16	2015/09/16	Live Birth	UNKNOWN	3690
53	101xxxxxxx	35	HIGH CONFIDENCE	✓	NEGATIVE (2011/12/22)	✓		2015/10/20	2015/10/20	2015/10/20		UNKNOWN	
54	148xxxxxxx	26	BIRTH OUTCOME	✓	NEGATIVE (2015/08/13)	✓		2015/06/15	2015/08/09	2015/08/09	Live Birth	UNKNOWN	3690
55	120xxxxxxx	24	BIRTH OUTCOME	✓	UNKNOWN	✓	TB positive	2015/02/26	2015/07/06	2015/07/06	Live Birth	UNKNOWN	2120
56	129xxxxxxx	30	BIRTH OUTCOME	✓	ART (2011/06/20)	✓	TB positive	2015/10/19	2016/03/23	2016/03/23	Live Birth	NEGATIVE(2016-06-01)	3220
57	260xxxxxxx	26	HIGH CONFIDENCE	✓	ART (2014/10/01)	✓	TB positive	2015/09/08	2015/09/08	2015/09/08		UNKNOWN	
58	142xxxxxxx	20	BIRTH OUTCOME	✓	UNKNOWN	✓		2015/07/13	2015/12/30	2015/12/30	Still Birth	UNKNOWN	2500

Sample output – HIV-infected pregnancies in 2015
 - Note that negatives are mostly not confirmed due to not having PoC test data

Row ID	Folder No	Age	Pregnancy Evidence	PIDA	HIV status	Dead	Co-morbidities	Pregnancy First Evidence Date	EDD	Pregnancy Outcome Date	Infant Outcome	Infant HIV Status	Birth Weight (g)
19873	142xxxxxxxx	26	BIRTH OUTCOME	✓	ART during (2015/08/04)			2015/12/22	2016/02/20	2016/02/20	Live Birth	NEGATIVE(2017-02-16)	3990
19874	142xxxxxxxx	25	BIRTH OUTCOME	✓	ART during (2015/12/22)			2015/12/22	2016/06/23	2016/06/23	Live Birth	NEGATIVE(2016-10-07)	3125
19875	258xxxxxxxx	29	BIRTH OUTCOME	✓	ART (2013/07/01)			2015/12/08	2016/04/24	2016/04/24	Live Birth	NEGATIVE(2016-07-05)	2560
19876	481xxxxxxxx	40	BIRTH OUTCOME	✓	ART during (2015/12/14)		Hypertension	2015/12/14	2016/03/23	2016/03/23	Live Birth, Live Birth	NEGATIVE(2016-11-10), NEGATIVE(2016-11-10)	1820, 1880
19877	120xxxxxxxx	26	BIRTH OUTCOME	✓	ART during (2015/12/17)			2015/12/17	2015/12/17	2015/12/17	Live Birth	NEGATIVE(2016-03-15)	3160
19878	893xxxxxxxx	38	BIRTH OUTCOME	✓	ART post natally (2016/04/27)			2015/12/22	2016/04/25	2016/04/25	Live Birth	NEGATIVE(2016-04-28)	2330
19879	140xxxxxxxx	36	HIGH CONFIDENCE	✓	ART (2005/08/10)			2015/12/22	2016/01/19				
19880	129xxxxxxxx	31	HIGH CONFIDENCE	✓	ART (2012/08/15)			2015/12/22	2016/04/12				
19881	112xxxxxxxx	26	HIGH CONFIDENCE	✓	ART (2012/11/23)		TB positive	2015/12/22	2016/04/12				
19882	114xxxxxxxx	19	BIRTH OUTCOME	✓	ART during (2015/06/15)			2015/12/21	2015/12/22	2015/12/22	Live Birth	NEGATIVE(2016-02-02)	2975
19883	135xxxxxxxx	30	BIRTH OUTCOME	✓	ART (2014/11/10)			2015/12/21	2016/05/07	2016/05/07	Live Birth	NEGATIVE(2016-07-21)	2780
19884	148xxxxxxxx	22	HIGH CONFIDENCE	✓	ART (2014/12/03)			2015/12/21	2015/12/21	2015/12/21		UNKNOWN	
19885	140xxxxxxxx	34	BIRTH OUTCOME	✓	ART (2008/10/17)			2015/12/21	2015/12/21	2015/12/21	Live Birth	UNKNOWN	3530
19886	114xxxxxxxx	24	HIGH CONFIDENCE	✓	ART during (2015/12/01)			2015/12/29	2016/01/26				
19887	155xxxxxxxx	28	HIGH CONFIDENCE	✓	POSITIVE (2011/07/04)			2015/12/27	2015/12/27	2015/12/27		UNKNOWN	
19888	441xxxxxxxx	36	HIGH CONFIDENCE	✓	ART (2009/11/12)		TB positive	2015/12/28	2016/01/25				
19889	254xxxxxxxx	32	HIGH CONFIDENCE	✓	ART during (2015/12/04)			2015/12/21	2016/01/18				
19890	260xxxxxxxx	35	BIRTH OUTCOME	✓	ART (2007/08/27)		Mental Health	2015/12/21	2016/01/23	2016/01/23	Live Birth	NEGATIVE(2016-01-23)	4530
19891	116xxxxxxxx	26	BIRTH OUTCOME	✓	ART (2011/02/01)		TB positive	2015/11/25	2016/03/11	2016/03/11	Live Birth	NEGATIVE(2016-06-28)	2820
19892	879xxxxxxxx	24	HIGH CONFIDENCE	✓	ART (2014/03/13)			2015/12/01	2016/03/22				
19893	154xxxxxxxx	23	BIRTH OUTCOME	✓	ART (2012/02/07)			2015/12/29	2016/01/06	2016/01/06		UNKNOWN	3200
19894	269xxxxxxxx	24	BIRTH OUTCOME	✓	ART (2010/10/11)		TB positive	2015/12/27	2015/12/27	2015/12/27	Live Birth	NEGATIVE(2016-02-10)	3030
19895	112xxxxxxxx	26	HIGH CONFIDENCE	✓	ART (2012/02/29)			2015/12/29	2016/01/26				
19896	827xxxxxxxx	30	HIGH CONFIDENCE	✓	ART (2013/07/25)		TB positive	2015/12/29	2016/04/19				
19897	132xxxxxxxx	31	BIRTH OUTCOME	✓	ART during (2015/09/09)			2015/12/22	2016/03/29	2016/03/29	Live Birth	NEGATIVE(2016-06-07)	3330
19898	460xxxxxxxx	25	BIRTH OUTCOME	✓	ART during (2015/12/01)			2015/12/01	2016/05/21	2016/05/21	Live Birth	NEGATIVE(2017-03-03)	2860
19899	106xxxxxxxx	30	BIRTH OUTCOME	✓	ART (2015/07/14)			2015/12/14	2016/06/04	2016/06/04	Live Birth	NEGATIVE(2016-08-16)	2690
19900	161xxxxxxxx	36	HIGH CONFIDENCE	✓	POSITIVE (2013/05/15)		Hypertension	2015/12/14	2016/04/04				
19901	226xxxxxxxx	46	HIGH CONFIDENCE	✓	ART during (2015/08/14)		Mental Health	2015/12/21	2016/01/18				
19902	258xxxxxxxx	33	BIRTH OUTCOME	✓	POSITIVE (2014/12/08)			2015/12/22	2016/05/04	2016/05/04	Live Birth	UNKNOWN	2670
19903	176xxxxxxxx	35	BIRTH OUTCOME	✓	ART (2008/09/26)			2015/12/22	2016/08/08	2016/08/08	Live Birth	NEGATIVE(2016-08-08)	3980

Sample output – HIV-infected births from 2015 pregnancies n=353, limited to high confidence

Health BI > Clinical BI > Data Centre > Reports > Confidential

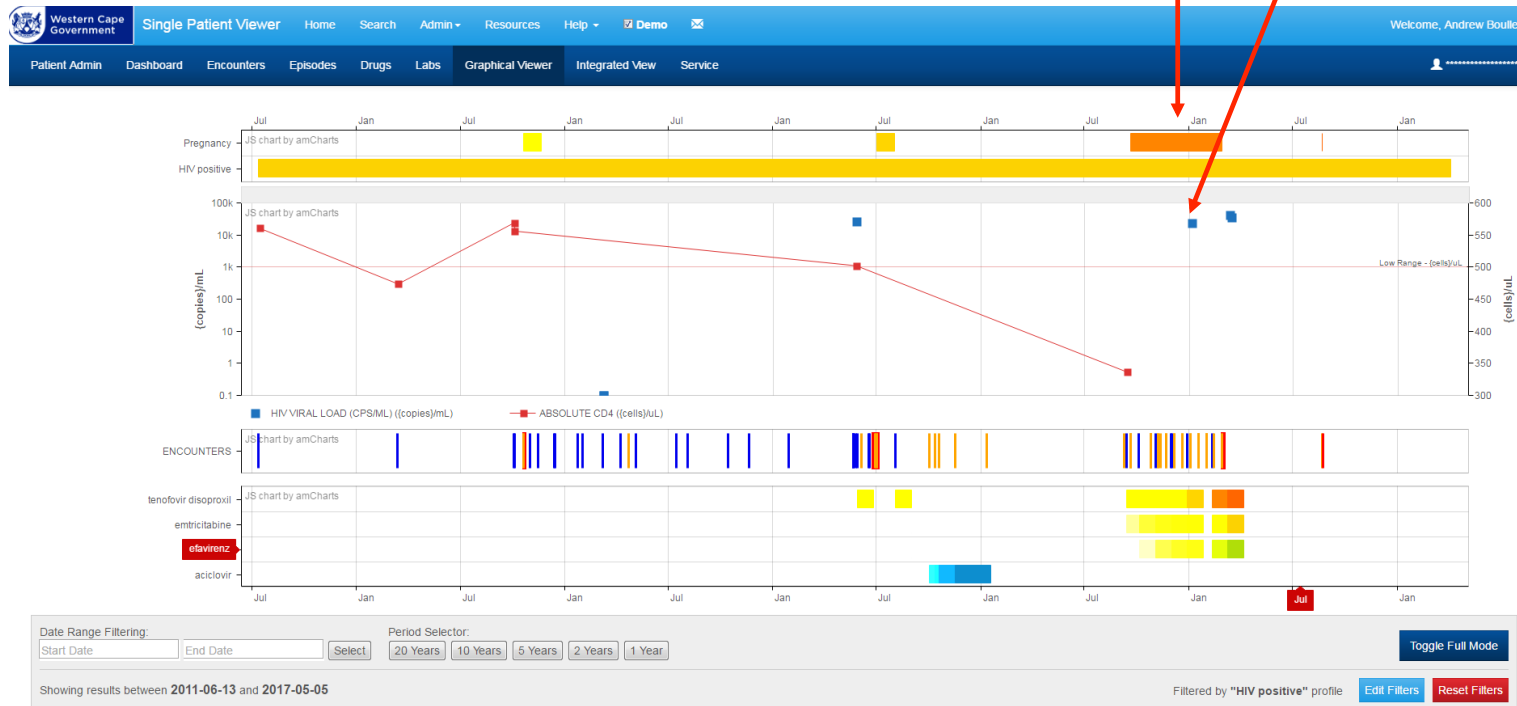
Actions 8 of 8 Find Next

Row ID	Folder No	Age	Pregnancy Evidence	Valid	HIV status	Dead	Co-morbidities	Pregnancy First Evidence Date	EDD	Pregnancy Outcome Date	Infant Outcome	Infant HIV Status	Birth Weight (g)
337	410xxxxxxx	35	BIRTH OUTCOME	✓	ART (2012/04/02)			2015/12/15	2015/12/16	2015/12/16	Live Birth	ART(2016-05-24)	2510
338	119xxxxxxx	25	BIRTH OUTCOME	✓	ART (2014/12/10)		TB positive	2015/12/04	2016/02/01	2016/02/01	Live Birth	POSITIVE(2016-03-17)	348
339	106xxxxxxx	35	BIRTH OUTCOME	✓	ART (2015/06/24)			2015/12/07	2016/06/17	2016/06/17	Live Birth	POSITIVE(2017-02-28)	3660
340	231xxxxxxx	26	BIRTH OUTCOME	✓	ART (2014/01/06)			2015/11/11	2016/04/01	2016/04/01	Live Birth	POSITIVE(2017-03-27)	2800
341	137xxxxxxx	22	BIRTH OUTCOME	✓	ART during (2015/11/30)			2015/11/23	2016/07/24	2016/07/24	Live Birth	ART(2016-10-26)	3340
342	377xxxxxxx	27	BIRTH OUTCOME	✓	ART (2015/03/27)		Hypertension, TB positive	2015/12/30	2016/03/02	2016/03/02	Live Birth	POSITIVE(2016-08-22)	2820
343	775xxxxxxx	24	BIRTH OUTCOME	✓	ART post nataly (2015/09/21)			2015/09/15	2015/09/16	2015/09/16	Live Birth	ART(2016-05-10)	3725
344	150xxxxxxx	22	BIRTH OUTCOME	✓	ART during (2015/05/15)			2015/11/06	2015/11/16	2015/11/16	Live Birth	POSITIVE(2016-02-01)	4270
345	118xxxxxxx	28	BIRTH OUTCOME	✓	ART (2011/12/07)		TB positive	2015/11/21	2015/11/21	2015/11/21	Live Birth	POSITIVE(2015-11-22)	2820
346	142xxxxxxx	21	BIRTH OUTCOME	✓	ART post nataly (2015/12/17)			2015/12/08	2015/12/08	2015/12/08	Live Birth	ART(2016-04-12)	3265
347	116xxxxxxx	20	BIRTH OUTCOME	✓	ART post nataly (2015/11/03)			2015/09/15	2015/09/15	2015/09/15	Live Birth	ART(2015-12-09)	3550
348	439xxxxxxx	31	BIRTH OUTCOME	✓	ART (2010/03/19)			2015/10/20	2015/10/20	2015/10/20	Live Birth	ART(2016-03-02)	2360
349	154xxxxxxx	33	BIRTH OUTCOME	✓	ART during (2015/11/02)			2015/10/29	2016/05/28	2016/05/28	Live Birth, Live Birth	POSITIVE(2016-05-31), POSITIVE(2016-05-31)	2740, 2990
350	407xxxxxxx	34	BIRTH OUTCOME	✓	ART during (2015/10/15)			2015/12/13	2015/12/13	2015/12/13	Live Birth	ART(2016-12-09)	2690
351	796xxxxxxx	23	BIRTH OUTCOME	✓	ART (2009/03/09)			2015/11/22	2015/11/22	2015/11/22	Live Birth	ART(2016-09-21)	2590
352	340xxxxxxx	22	BIRTH OUTCOME	✓	ART during (2016/07/28)			2015/12/15	2016/08/02	2016/08/02	Live Birth	ART(2016-08-15)	2710
353	451xxxxxxx	24	BIRTH OUTCOME	✓	ART (2012/05/24)			2015/10/12	2016/01/22	2016/01/22	Live Birth	POSITIVE(2016-02-15)	1030

Hot link to single patient viewer on a patient with an HIV positive transmission from a 2015 pregnancy

Pregnancy with HIV outcome
3rd pregnancy of 4, last one was
terminated in 2016

Viral load elevated 2m
before delivery



Folder #	ID No	First Name	Last Name	Address	Date of Birth	Date of Death	Age	Gender	Link Patients	Group Episodes
713044253	7905257130442	NHLAKANIPO CLIVE	KIMBUNGU	Street X, Suburb X, City X, Western Cape	1979-05-25		37	F		

Folder #	Episode	Episode Date	Treatment Start Date	Last Contact Date
713044253	Diabetes Mellitus	2013-11-01	2013-11-01	2016-12-21
713044253	Hypertension	2013-04-18	2013-04-18	2016-12-21
713044253	Pregnancy	2000-08-13	2000-08-13	2000-08-13

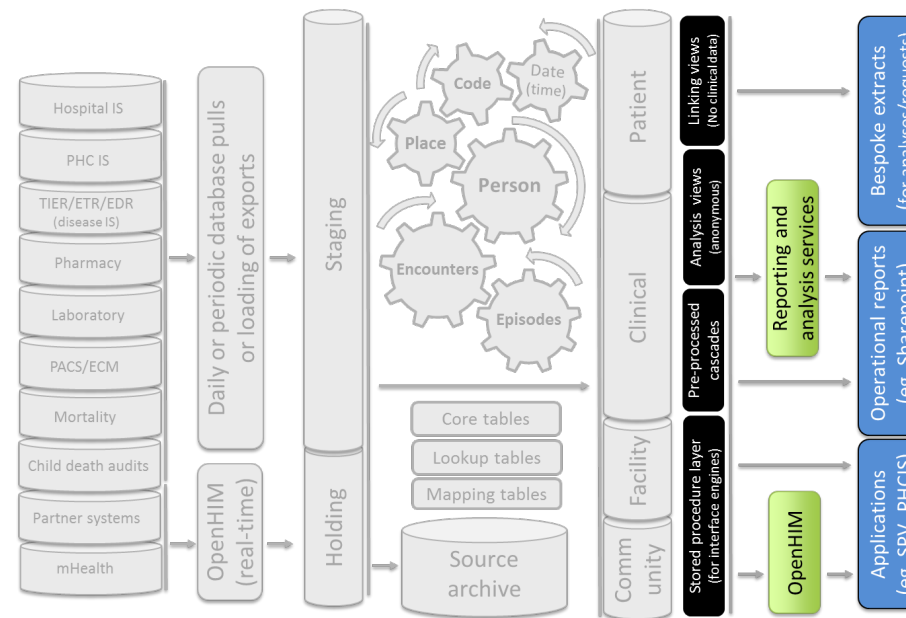
Folder #	Date	Facility	Lab Description	Result
713044253	2017-02-04	Facility 344	GLYCATED HAEMOGLOBIN (NGSP)	6.1
713044253	2017-02-04	Facility 344	GLYCATED HAEMOGLOBIN (IFCC)	43
713044253	2017-02-04	Facility 344	EST. AVERAGE GLUCOSE (EAG)	7.1
713044253	2017-02-04	Facility 344	GLYCATED HAEMOGLOBIN AUTO COM... MONITORING GLYCAEMIC CONT...	
713044253	2017-02-04	Facility 344	CREATININE	77
713044253	2017-02-04	Facility 344	MDRD FORMULA	>60
713044253	2017-02-04	Facility 344	TOTAL CHOLESTEROL	3.69
713044253	2014-05-23	Facility 344	GLYCATED HAEMOGLOBIN (NGSP)	6.4
713044253	2014-03-01	Facility 319	ALKALINE PHOSPHATASE (ALP)	220
713044253	2014-03-01	Facility 319	HAEMOGLOBIN	11.3

Folder #	Visit Date	Discharge Date	Facility	Type	ICD	Links
713044253	2017-02-03		Facility 344			
713044253	2016-05-17		Facility 344			
713044253	2016-04-14		Facility 344			
713044253	2014-05-22		Facility 344			
713044253	2014-03-06		Facility 319			
713044253	2014-03-03		Facility 319			
713044253	2014-02-24		Facility 319			
713044253	2014-02-14		Facility 319			
713044253	2014-02-10		Facility 319			
713044253	2014-02-09	2014-03-03	Facility 319		Cholecystitis, unspecified	

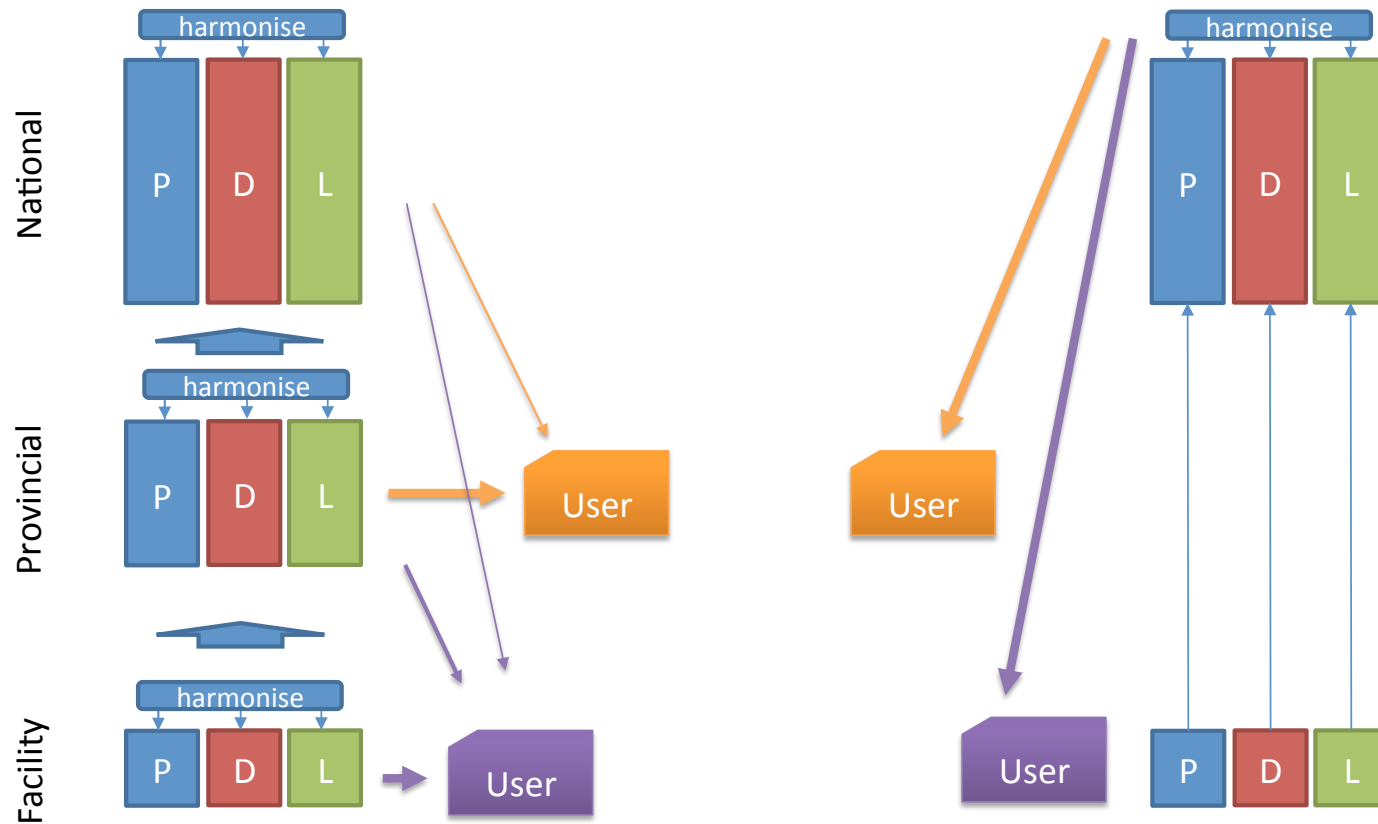
Folder #	Date	Facility	Drug	Quantity	Source
713044253	2017-02-03	Facility 344	LANSOPRAZOLE	10 Capsule	JAC
713044253	2017-02-03	Facility 344	ISOSORBIDE DINITRATE	1 x 50 Tablet Pack	JAC

Outline

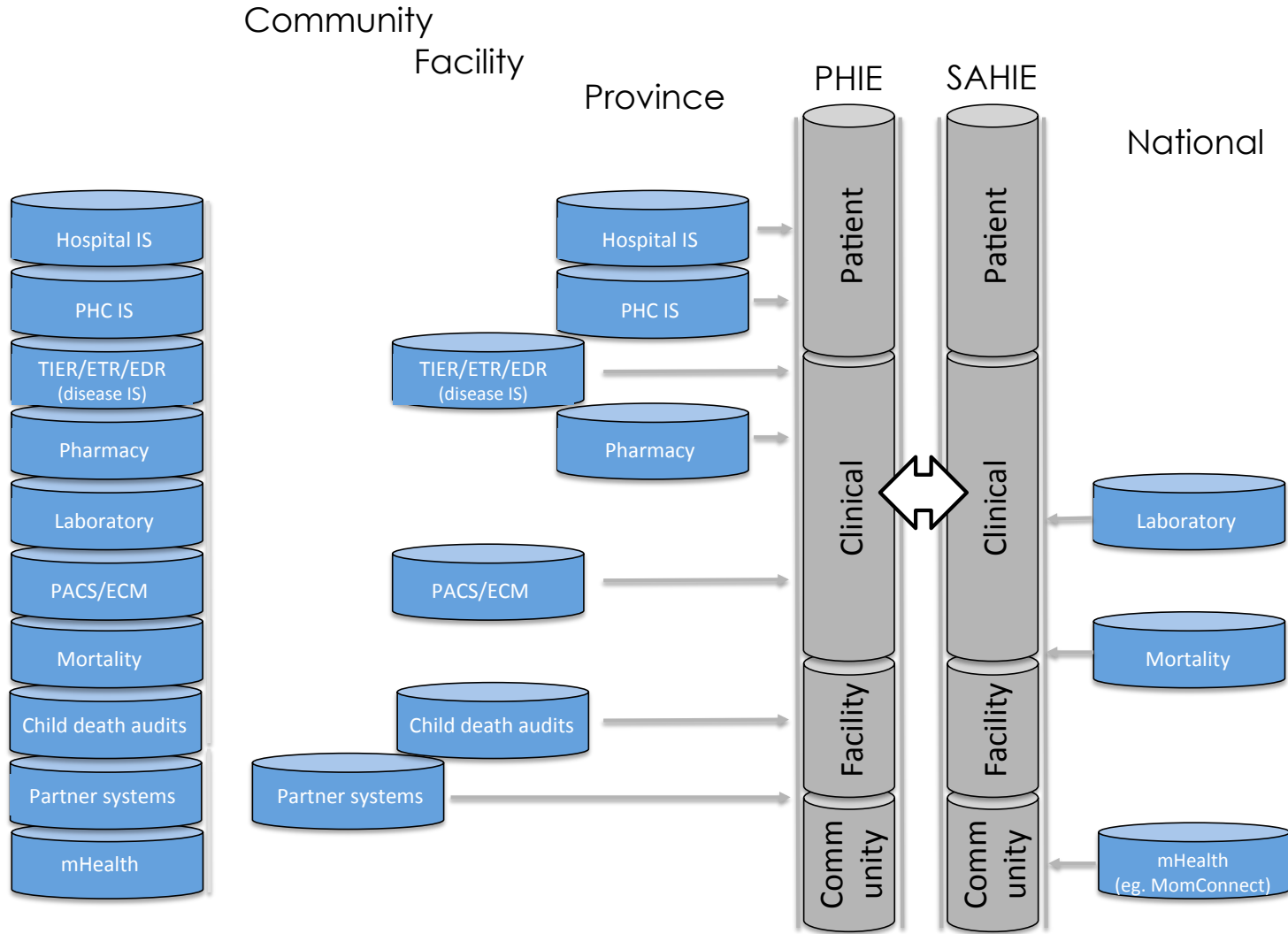
- ❑ Introduction, background, architecture
- ❑ Data curation
- ❑ Data beneficiation and enrichment: encounters, episodes and cascades
- ❑ Reporting
- ❑ **Data governance**



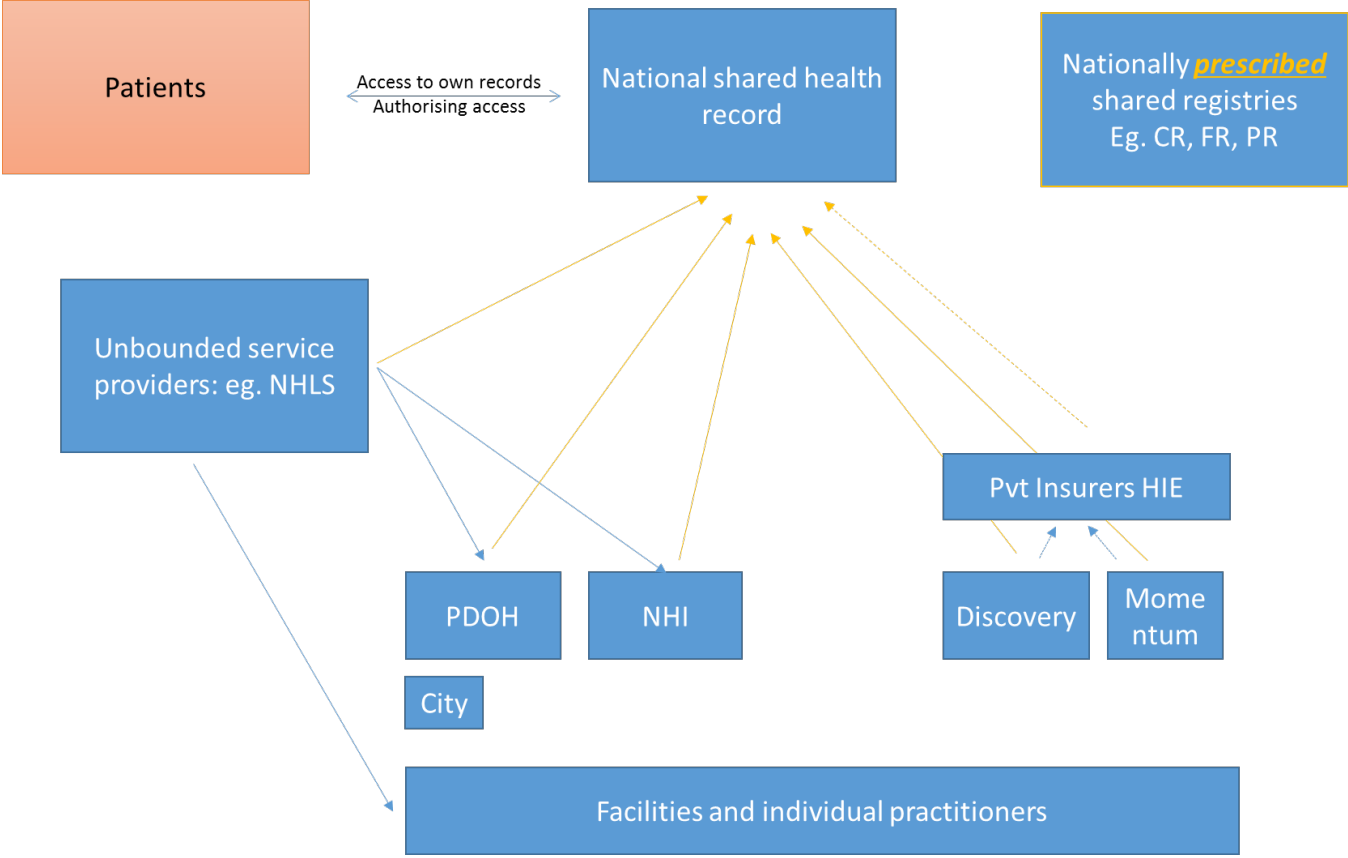
Architectures for facility/province/national data harmonisation



Architectures for facility/province/national data harmonisation



Role of a Provincial Health Data Centre



Immediate opportunities for public-private data sharing

- Trusted third party could link to understand cross-sector utilisation
 - On ID number
 - On encrypted ID number
 - On all identifiers
 - On encrypted versions of all identifiers

Schmidlin *et al.* *BMC Medical Research Methodology* (2015) 15:46
DOI 10.1186/s12874-015-0038-6



RESEARCH ARTICLE

Open Access

Privacy Preserving Probabilistic Record Linkage (P3RL): a novel method for linking existing health-related data and maintaining participant confidentiality



Kurt Schmidlin¹, Kerri M. Clough-Gorr^{1,2}, and Adrian Spoerri^{1*} for the SNC study group

Please note

This presentation is incomplete
without the accompanying narrative