

Figure 1: Example admitted episodes data structure.

	id	admdate	sepdate	admmode	sepmode	diag1	diag2
1.	1	01jan2020	02jan2020	H	T	I214	R073
2.	1	02jan2020	03jan2020	T	T	I214	I10
3.	1	03jan2020	09jan2020	T	H	I214	R073
4.	1	20apr2021	25apr2021	H	H	I509	R060
5.	2	16aug2018	03sep2018	H	H	I639	R418
6.	2	03mar2019	03mar2019	H	H	D649	E119
7.	2	03mar2019	03mar2019	H	H	D649	E119
8.	2	24nov2021	26nov2021	H	H	N179	E119
9.	2	30dec2022	02jan2023	H	H	K922	E119
10.	2	11may2023	15may2023	H	D	J189	E119

- *id* – ID number for an individual
- *admdate* – admission date
- *sepdate* – separation date
- *admmode* – admission mode, where H is home and T is transfer
- *sepmode* – separation mode, where H is home, T is transfer, and D is death
- *diag1* – diagnosis 1 or primary diagnosis (ICD-10 code)
- *diag2* – diagnosis 2 or secondary diagnosis (ICD-10 code). There are more secondary diagnoses.

Figure 2: Common aspects of hospital admissions data requiring processing when defining episodes of care.

A - Duplicate admissions

	id	admdate	sepdate	admcode	sepmode	diag1
1.	1	01jan2020	01jan2020	H	H	I214
2.	1	01jan2020	01jan2020	H	H	I214
3.	1	01jan2020	01jan2020	H	D	J189

B - Nested admissions

	id	admdate	sepdate	diag1	diag2
1.	1	01jan2020	08jan2020	I214	
2.	1	02jan2020	02jan2020	I214	E119
3.	1	03jan2020	03jan2020	I499	E119
4.	1	05jan2020	05jan2020	J152	E119
5.	1	06jan2020	12jan2020	I214	E119
6.	1	15jul2020	15jul2020	I509	E119

C - Transfers

	id	admdate	sepdate	diag1	diag2	admcode	sepmode
1.	1	01jan2020	01jan2020	I214		H	T
2.	1	02jan2020	03jan2020	I214	E119	T	T
3.	1	03jan2020	05jan2020	I499	E119	T	H
4.	1	05jan2020	06jan2020	J152	E119	T	T
5.	1	06jan2020	10jan2020	I214	E119	H	T
6.	1	10jan2020	15jan2020	I214	E119	T	H
7.	1	15jul2020	15jul2020	I509	E119	T	D

Figure 3: Processing duplicate admissions.

	id	admdate	sepdate	diag1	MI	DM
1.	1	01jan2020	01jan2020	I214	1	.
2.	1	01jan2020	01jan2020	E119	.	1
3.	1	01jan2020	01jan2020	I214	1	.

↓ 1: tag duplicates

	id	admdate	sepdate	diag1	MI	DM	dup
1.	1	01jan2020	01jan2020	I214	1	.	.
2.	1	01jan2020	01jan2020	E119	.	1	1
3.	1	01jan2020	01jan2020	I214	1	.	1

↓ 2: de-tag duplicates that aren't first

	id	admdate	sepdate	diag1	MI	DM	dup
1.	1	01jan2020	01jan2020	I214	1	.	.
2.	1	01jan2020	01jan2020	E119	.	1	1
3.	1	01jan2020	01jan2020	I214	1	.	.

↓ 3: collect data from the second admission

	id	admdate	sepdate	diag1	MI	DM	dup
1.	1	01jan2020	01jan2020	I214	1	1	.
2.	1	01jan2020	01jan2020	E119	.	1	1
3.	1	01jan2020	01jan2020	I214	1	.	.

↓ 4: drop the second admission

	id	admdate	sepdate	diag1	MI	DM
1.	1	01jan2020	01jan2020	I214	1	1
2.	1	01jan2020	01jan2020	I214	1	.

↓ Repeat

	id	admdate	sepdate	MI	DM
1.	1	01jan2020	01jan2020	1	1

Figure 4: Processing nested admissions.

	id	admdate	sepdate	diag1	diag2	MI	DM	AR
1.	1	01jan2020	08jan2020	I214		1	.	.
2.	1	02jan2020	02jan2020	I214	E119	1	1	.
3.	1	03jan2020	03jan2020	I499	E119	.	1	1
4.	1	05jan2020	05jan2020	J152	E119	.	1	.
5.	1	06jan2020	12jan2020	I214	E119	1	1	.

↓ 1: tag nested admissions

	id	admdate	sepdate	diag1	diag2	MI	DM	AR	nest
1.	1	01jan2020	08jan2020	I214		1	.	.	.
2.	1	02jan2020	02jan2020	I214	E119	1	1	.	1
3.	1	03jan2020	03jan2020	I499	E119	.	1	1	.
4.	1	05jan2020	05jan2020	J152	E119	.	1	.	.
5.	1	06jan2020	12jan2020	I214	E119	1	1	.	.

↓ 2: de-tag nested admissions that' aren't first

	id	admdate	sepdate	diag1	diag2	MI	DM	AR	nest
1.	1	01jan2020	08jan2020	I214		1	.	.	.
2.	1	02jan2020	02jan2020	I214	E119	1	1	.	1
3.	1	03jan2020	03jan2020	I499	E119	.	1	1	.
4.	1	05jan2020	05jan2020	J152	E119	.	1	.	.
5.	1	06jan2020	12jan2020	I214	E119	1	1	.	.

↓ 3: collect data from the second admission

	id	admdate	sepdate	diag1	diag2	MI	DM	AR	nest
1.	1	01jan2020	08jan2020	I214		1	1	.	.
2.	1	02jan2020	02jan2020	I214	E119	1	1	.	1
3.	1	03jan2020	03jan2020	I499	E119	.	1	1	.
4.	1	05jan2020	05jan2020	J152	E119	.	1	.	.
5.	1	06jan2020	12jan2020	I214	E119	1	1	.	.

↓ 4: drop the second admission

	id	admdate	sepdate	diag1	diag2	MI	DM	AR
1.	1	01jan2020	08jan2020	I214		1	1	.
2.	1	03jan2020	03jan2020	I499	E119	.	1	1
3.	1	05jan2020	05jan2020	J152	E119	.	1	.
4.	1	06jan2020	12jan2020	I214	E119	1	1	.

↓ Repeat

	id	admdate	septime	diag1	diag2	MI	DM	AR
1.	1	01jan2020	08jan2020	I214		1	1	1
2.	1	05jan2020	05jan2020	J152	E119	.	1	.
3.	1	06jan2020	12jan2020	I214	E119	1	1	.

↓ Repeat

	id	admdate	septime	diag1	diag2	MI	DM	AR
1.	1	01jan2020	08jan2020	I214		1	1	1
2.	1	06jan2020	12jan2020	I214	E119	1	1	.

↓ Repeat

	id	admdate	septime	diag1	diag2	MI	DM	AR
1.	1	01jan2020	12jan2020	I214		1	1	1

↓ Repeat

	id	admdate	septime	MI	DM	AR
1.	1	01jan2020	12jan2020	1	1	1

Figure 5: Processing transfers.

	id	admdate	sepdate	diag1	diag2	admmode	sepmode	MI
1.	1	01jan2020	02jan2020	I214		H	T	1
2.	1	02jan2020	03jan2020	I214	E119	T	T	1
3.	1	03jan2020	05jan2020	I499	E119	T	H	.
4.	1	05jan2020	06jan2020	J152	E119	T	T	.
5.	1	06jan2020	10jan2020	I214	E119	T	T	1
6.	1	10jan2020	15jan2020	I214	E119	T	H	1

↓ 1: tag potential transfers admissions

	id	admdate	sepdate	diag1	diag2	admmode	sepmode	MI	ptr
1.	1	01jan2020	02jan2020	I214		H	T	1	.
2.	1	02jan2020	03jan2020	I214	E119	T	T	1	1
3.	1	03jan2020	05jan2020	I499	E119	T	H	.	1
4.	1	05jan2020	06jan2020	J152	E119	T	T	.	1
5.	1	06jan2020	10jan2020	I214	E119	T	T	1	1
6.	1	10jan2020	15jan2020	I214	E119	T	H	1	1

↓ 2: confirm based on admission and separation dates

	id	admdate	sepdate	diag1	diag2	admmode	sepmode	MI	ptr	tr
1.	1	01jan2020	02jan2020	I214		H	T	1	.	.
2.	1	02jan2020	03jan2020	I214	E119	T	T	1	1	1
3.	1	03jan2020	05jan2020	I499	E119	T	H	.	1	1
4.	1	05jan2020	06jan2020	J152	E119	T	T	.	1	1
5.	1	06jan2020	10jan2020	I214	E119	T	T	1	1	1
6.	1	10jan2020	15jan2020	I214	E119	T	H	1	1	1

↓ 3: de-tag transfers that aren't first in a set

	id	admdate	sepdate	diag1	diag2	admmode	sepmode	MI	ptr	tr
1.	1	01jan2020	02jan2020	I214		H	T	1	.	.
2.	1	02jan2020	03jan2020	I214	E119	T	T	1	1	1
3.	1	03jan2020	05jan2020	I499	E119	T	H	.	1	.
4.	1	05jan2020	06jan2020	J152	E119	T	T	.	1	1
5.	1	06jan2020	10jan2020	I214	E119	T	T	1	1	.
6.	1	10jan2020	15jan2020	I214	E119	T	H	1	1	1

↓ 4: collect data from the second admission

	id	admdate	sepdate	diag1	diag2	admmode	sepmode	MI	ptr	tr
1.	1	01jan2020	03jan2020	I214		H	T	1	.	.
2.	1	02jan2020	03jan2020	I214	E119	T	T	1	1	1
3.	1	03jan2020	06jan2020	I499	E119	T	T	.	1	.
4.	1	05jan2020	06jan2020	J152	E119	T	T	.	1	1
5.	1	06jan2020	15jan2020	I214	E119	T	H	1	1	.
6.	1	10jan2020	15jan2020	I214	E119	T	H	1	1	1

↓ 5: drop the second admission

	id	admdate	septime	diag1	diag2	admmode	septime	MI
1.	1	01jan2020	03jan2020	I214		H	T	1
2.	1	03jan2020	06jan2020	I499	E119	T	T	.
3.	1	06jan2020	15jan2020	I214	E119	T	H	1

↓ Repeat

	id	admdate	septime	diag1	diag2	admmode	septime	MI
1.	1	01jan2020	06jan2020	I214		H	T	1
2.	1	06jan2020	15jan2020	I214	E119	T	H	1

↓ Repeat

	id	admdate	septime	admmode	septime	MI
1.	1	01jan2020	15jan2020	H	H	1

For *admmode* and *septime*, H=Home and T=Transfer.