Cost centre	ID	Dispense date	Product	Dispense net pack quantity
GP Clinic	1	20/02/2018	Abacavir-Lamivudine	4 packs
GP Clinic	1	20/02/2018	Darunavir-Cobicistat	4 packs
GP Clinic	1	20/02/2018	Dolutegravir	4 packs
GP Clinic	2	20/06/2018	Elvitegravir-Cobicistat-Emtricitabine-Tenofovir Alafenamide	2 packs
GP Clinic	2	13/08/2018	Elvitegravir-Cobicistat-Emtricitabine-Tenofovir Alafenamide	2 packs
GP Clinic	2	13/08/2018	Elvitegravir-Cobicistat-Emtricitabine-Tenofovir Alafenamide	2 packs

↓

	ID	Dispense date	Product	Dispense net pack quantity
GP Clinic	1	20/02/2018	Abacavir-Lamivudine Darunavir-Cobicistat Dolutegravir	4 packs
GP Clinic	2	20/06/2018	Elvitegravir-Cobicistat-Emtricitabine-Tenofovir Alafenamide	2 packs
GP Clinic	2	13/08/2018	Elvitegravir-Cobicistat-Emtricitabine-Tenofovir Alafenamide	4 packs

Figure S1. An example of the data cleaning process demonstrating combining multiple drugs recorded on the same date. If the drugs are different, they are combined as a single combination therapy. If the drugs are the same, they are combined as a single dispensing record and the dispense net pack quantity is multiplied.