# Algorithm for defining hospital stays

Katrine Damgaard, Jon Helgeland
Norwegian Knowledge Centre for the Health Services, Oslo, Norway

## Objective

To study outcomes of hospital care, such as survival within 30 days or readmissions, data must be organized into manageable hospital stays or episode of care. Hospitalized patients are often transferred between wards, departments and hospitals.  Our present objective is to develop an algorithm capturing the patients' complete admission histories, by linking their separate hospital and/or ward admissions into one single admissions chain (episode of care).

## Material and methods

Tomislav Dimoski at the Norwegian Knowledge Centre for the Health Services has developed a system for retrieving data from the patient administrative systems (PAS) at all Norwegian hospitals. By use of this system, we retrieved PAS data at each hospital for patients discharged during 2005-2009.

Each record corresponds to one admission at one ward within a department in a hospital. To define a complete patient stay, we have to aggregate the admissions; department stays, hospital stays and stays involving more than one hospital. All permanent residents in Norway have a personal identification number (PIN) which enables linking the records within and between hospitals. The data are not necessarily consistent: times may not be accurately recorded and seemingly overlapping admissions occur.

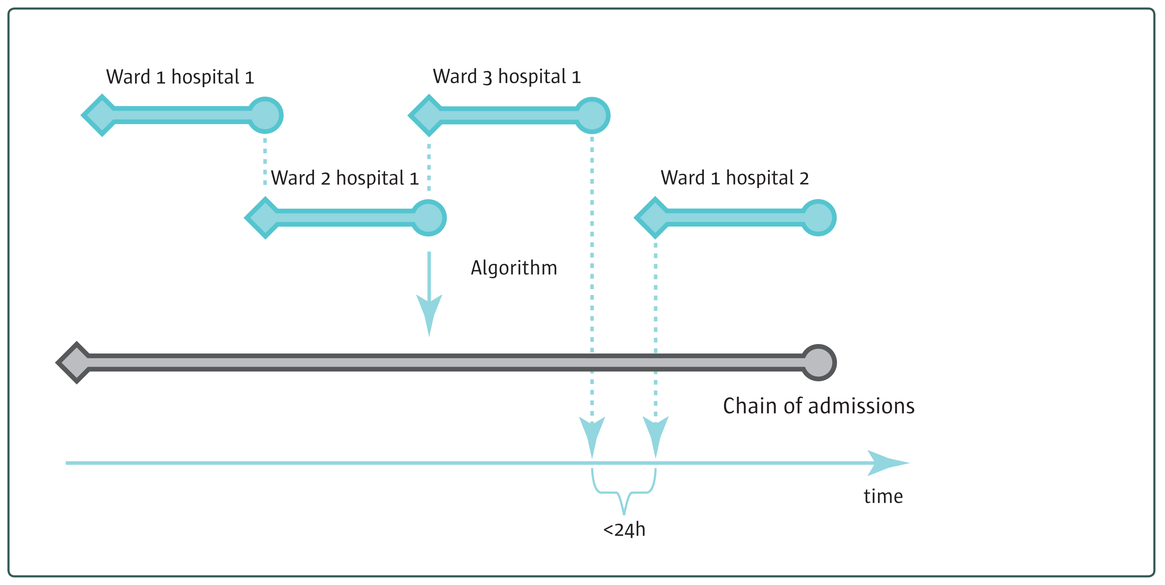
We have developed and tested an algorithm for concatenating the ward admissions to a chain of admissions, giving all the levels of stays described above. The input is a serial number, PIN, date of admission, date of discharge, hospital, department and further optional parameters.  The function uses the package **multicore** in R, for its ability to use more than one processor for large datasets. One of the properties of the algorithm is that you can choose different time tolerances for concatenating the ward admissions: if the interval between admissions exceeds the tolerance limit it will start a new hospital stay. We chose 24 hours as tolerance for our purpose.

## Results

With the choice of 24 hours time tolerance for difference between time of discharge and next admission, we found 10 485 022 hospital stays involving one or more hospitals out of 16 370 163 ward admissions from 3 304 546 patients. The running time of the algorithm is approximately two hours on a six-processor server.

Figure 1  How the algorithm links the ward admissions

A. An ideal situation




B. A moderately complex situation

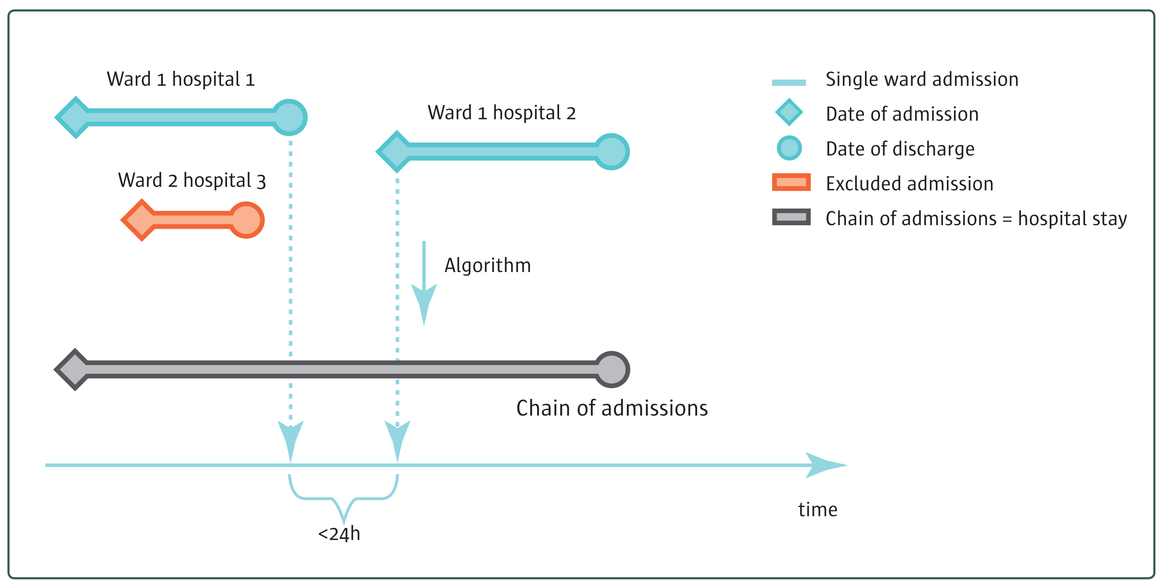



Figure 1 shows two examples of how the algorithm links the ward admissions to one complete patient hospital stay. The first situation, A, is the ideal one, where the date and time of discharge exactly matches the next admission. It also shows a transfer to another hospital within the accepted time (in this example 24 hours). The second example, B, shows a more unclear situation where we find a new ward admission during another admission. This may be due to e.g. Ward 1 forgot to register transfer out to and in from Ward 2 or that the stay at Ward 2 is wrongly allocated to this patient. This algorithm excludes the ward admission within another admission and thus makes is possible to handle such unclear data by excluding such unrealistic admissions.

## Conclusion

The algorithm can be used to aggregate large patient administrative data sets, with acceptable running times. The tolerance limit can be adjusted to suit the purpose and the data at hand. It handles the inconsistent data that are seen in practice.