Revisiting the incidence of schizophrenia: learning about the other half
Hogerzeil, S.J.

6

Design choices

Letter to the editor

Anderson et al. (2018) estimate the gap between the number of incident cases of schizophrenia aged 16–50 in Ontario, Canada in 1997–2015, and the number who were enrolled into Early Psychosis Intervention (EPI) services. Their analysis is a direct comparison between administrative records and the standard method for estimating the treated incidence.

The standard method (known as the 'first contact design') involves screening subjects for signs of psychosis when they present for psychiatric treatment. Subjects screened positive then undergo standardized diagnostic procedures to establish the criteria for schizophrenia. But studies based on administrative records have suggested that two out of three cases may be missed this way (Hogerzeil et al., 2014; Jongsma et al., 2019; Pedersen et al., 2014).

Anderson et al. (2018) also found that two out of three cases of schizophrenia had remained unknown to EPI services. Not all cases met EPI-services' inclusion criteria, but still a substantial number of true cases of schizophrenia had been missed. In their discussion, Anderson et al. (2018) focused on the issue of incomplete coverage of services, but this is only one of several design aspects that matter.
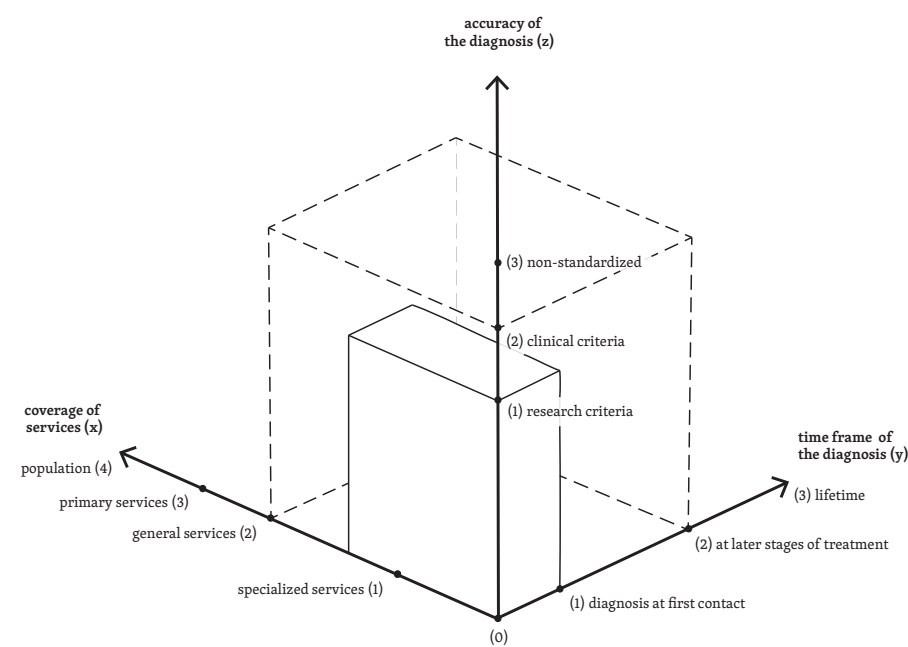
We propose to distinguish three design aspects where complete case finding can go wrong: coverage of services, time frame of the diagnosis, and accuracy of the diagnosis. We believe that these distinctions can help to understand the five- to ten-fold variation in incidence between populations, which is commonly reported but only partially explained (Jongsma et al., 2019, 2018; McGrath et al., 2004).

Coverage of services where cases can be detected. These may range from (1) very specialized services such as EPI services, emergency or inpatient services, extending to (2) general psychiatric or addiction services, and further to (3) primary care or somatic medical care or ultimately to (4) the general population.

Time frame of the diagnosis, the interval allowed between the first contact with a service and the moment a diagnosis can be made. It may range from (1) case ascertainment at first contact only, extending to (2) later stages of treatment, e.g. subjects presenting initially with another diagnosis, ultimately extending to a (3) life-time follow-up.

Accuracy of the diagnosis, ranging from diagnosis based on (1) research diagnostic pro-
cedures, extending to (2) clinical criteria diagnoses (e.g. DSM-5 or ICD-10) and (3) non
standardized diagnostic procedures.

This can be illustrated in 3D, where design choices along the x y z axes determine a box,
the volume of which represents the incidence estimate. Figure 6.1 illustrates how the first

accuracy of
the diagnosis (z)

(3) non-standardized

(2) clinical criteria

(1) research criteria

coverage of
services (x)

population (4)

primary services (3)

general services (2)

specialized services (1)

time frame of
the diagnosis (y)

(3) lifetime

(2) at later stages of treatment

(1) diagnosis at first contact

(0)

Figure 6.1 Graphical illustration of three design aspects for studies measuring the incidence of schizo-
phrenia
*The volume of the solid box represents the incidence as estimated in a typical first- contact design and that of the dotted box the
incidence as estimated in electronic administrative records.*

contact design (solid box; i.e. typically measured as first contacts at specialized services,
using research diagnoses) results in a lower incidence compared to cumulative records
(dotted box; i.e. typically measured at all psychiatric services, using clinical diagnoses
over much longer time spans).

Case-register studies from the 1950s to the 1970s typically focused on inpatient hospital
services, with long time frames and non-standardized diagnoses. The first contact stud-
ies of the 1990s and later focused on a wider coverage of services and better diagnostic
accuracy, while restricting the time frame (Jablensky et al., 1992). That approach has high
specificity but low sensitivity: many subjects with an ultimate diagnosis of schizophre-

nia will be missed because they do not meet criteria for the disorder when they first seek treatment (Hogerzeil et al., 2014; Rietdijk et al., 2011).

Longer time frames became possible with (a) the wide adoption DSM or ICD based clinical diagnoses and (b) well maintained administrative records in (c) institutions serving all psychiatric needs of well-defined populations. Such databases can now be used to reconstruct diagnostic histories or treatment pathways through services, up to the first diagnosis of schizophrenia, capturing new onsets along pathways that cannot be covered with the standard approach. This new approach is more sensitive, although it might come at the expense of some diagnostic specificity.

The study by Anderson et al. (2018) is the second to compare first contact and cumulative methods directly. Their study can be understood as a replication of our finding (Hogerzeil et al., 2014) that in administrative data the incidence of treated schizophrenia is two to threefold higher than detected using the first contact design. Now replicated, this finding has obvious implications for estimates of the number of cases affected, and for the organization of services. Furthermore, considering subjects with psychosis at first contact as 'prototype cases' may have distorted our understanding of schizophrenia by spuriously highlighting a younger age of onset and a more acute clinical presentation than seen in actual administrative records (Hogerzeil et al., 2016).

Study design matters a lot when estimating the incidence of schizophrenia. To interpret incidence studies or to make meaningful comparisons between them, we need a more elaborate classification of study designs, as suggested here.