



Universiteit
Leiden
The Netherlands

Revisiting the incidence of schizophrenia: learning about the other half

Hogerzeil, S.J.

Citation

Hogerzeil, S. J. (2022, March 10). *Revisiting the incidence of schizophrenia: learning about the other half*. Retrieved from <https://hdl.handle.net/1887/3278776>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3278776>

Note: To cite this publication please use the final published version (if applicable).

7

Impact of differences in study design

Published as:

Hogerzeil, S. J., Hoek, H. W., & Hemert, A. M. van. (2021). The impact of study design on schizophrenia incidence estimates: A systematic review of Northern European studies 2008-2019. *Schizophrenia Research*, 231, 134-141. <https://doi.org/10.1016/j.schres.2021.03.017>

Abstract

BACKGROUND — The best estimates of the incidence of schizophrenia range more than 25-fold from 3 to 80 per 100 000 person years. To what extent do differences in study design explain this wide variation?

METHOD — We selected all studies published between 2008–2019 reporting the incidence of schizophrenia in general populations of Northern Europe. We identified 17 estimates covering 85 million person years and more than 15 000 individual cases. The estimates ranged from 4–72 per 100 000 person years (median 30; interquartile range 13–41). We classified the estimates in terms of three study design factors (coverage of services, time frame, and diagnostic quality) and two population factors (urbanicity and age).

RESULTS — A meta-regression model of the three design factors, using the two population factors as covariates, explained 91% of between-study variation. Studies performed in general psychiatric services reported similar estimates [incidence rate ratio 1.12 (95% confidence interval 0.88 to 1.43)] to those performed in specialized services. But studies applying a cumulative time frame to diagnosis reported fourfold higher estimates [4.04 (3.14 to 5.2)] than those applying a first contact time frame. And studies based on clinical diagnoses reported lower estimates [0.55 (0.43 to 0.72)] than those based on standardized research diagnoses. The three study design factors by themselves explained 67% of between-study variation.

CONCLUSION — When comparing incidence rates from different populations, distortions arising from differences in study design can eclipse differences caused by schizophrenia risk factors, such as gender, age or migrant status.

Introduction

Rationale

Systematic reviews report a wide variation between estimates of the incidence of schizophrenia. Two international reviews together cover the period 1950–2017: one review of schizophrenia incidence studies published between 1950–2000 reported estimates ranging from 4–52 per 100 000 person years (van der Werf et al., 2012), while the other review of psychosis incidence studies published between 2002–2017 reported schizophrenia incidence estimates ranging from 3–76 per 100 000 person years (Jongsma et al., 2019). Variation between countries with different cultures and health care systems can be expected, but reviews of incidence from similar countries also show wide variations: a review of UK studies published in 1950–2009 reported estimates ranging from 4–32 per 100 000 person years (Kirkbride, Jones, et al., 2012); while another review of studies published between 1992–2012 with estimates from the Netherlands, Sweden and Denmark ranged from 9–80 per 100 000 person years (Vassos et al., 2012).

One explanation for the wide variation is that different rates result from different population characteristics, i.e. with different distributions of risk factors for schizophrenia. Populations with higher numbers of young adults or higher numbers of males, for example, are likely to report higher incidences than studies focusing on the population at large (Jongsma & Jones, 2019; Thorup et al., 2007). Similarly, studies in larger cities commonly report higher incidences than studies from rural areas (Vassos et al., 2016), and rates tend to be higher among immigrants than native inhabitants in an area (Bourque et al., 2010; Selten et al., 2019).

Another explanation for the variation could be that different rates result from different study designs. In a previous study, we used two different study designs to estimate the incidence of schizophrenia in the city of The Hague in the Netherlands (Hogerzeil et al., 2014). The first approach we used was a standard first contact design, which is generally considered the standard for incidence studies of schizophrenia. The second approach was based on a longitudinal case-register extracted from electronic hospital records. In the database, we could follow patients beyond their first contact to detect diagnostic changes over the course of treatment. This longitudinal case-register approach resulted in an estimate that was more than three times higher than the estimate based on the first contact approach [69 (95% confidence interval (CI) 64 to 74) vs. versus 21 (18 to 23) per 100 000 person years]. The impact of single aspects of study design was explored in several world-wide meta-analyses that included studies from heterogeneous populations (Bourque et al., 2010; McGrath et al., 2004; van der Werf et al., 2012). These analyses

uncovered no clear patterns. Two recent meta-analyses (Castillejos et al., 2019; Jongasma et al., 2019) examined this issue using meta-regression. Castillejos et al. (2019) reviewed only the literature based on first contact sampling and reported that methodological differences helped to explain between-study heterogeneity. Jongasma et al. (2019) compared case registers with first contact studies and reported that register-based estimates are systematically higher [with a multivariable model relative risk of 2.51 (95% CI 1.24 to 5.21)]. However, neither review was set up to quantify the relative importance of different factors in study design.

Objectives

We have previously proposed to categorize the design of incidence studies on three factors: *coverage of services*, *time frame of the diagnosis*, and *reliability of the diagnosis* (Hogerzeil & Hemert, 2019).

Our aim in this review was to examine to what extent reported incidence estimates are related to these three design factors, and so to distinguish artifacts from ‘true’ variation due to population characteristics. We hypothesized that estimates would be higher in studies with a wider service cover, longer time frames, and clinically oriented diagnoses.

To test this, we systematically identified all studies on the incidence of schizophrenia published from 2008–2019. We used meta-regression analysis to examine the impact of design features on the incidence estimates, adjusting for the impact of population characteristics.

Methods

This meta-analysis and meta-regression followed PRISMA guidelines (Liberati et al., 2009).

We based our study on the recent meta-analysis by Jongasma et al. (2019), which covered all the original research on the incidence of non-organic, adult-onset psychotic disorder published in 2002–2017. Her method in turn was based on a previous systematic review by Kirkbride, Errazuriz, et al. (2012), which covered the research conducted in England on the incidence of non-organic adult-onset psychosis, published in 1950–2009. Jongasma et al. (2019)’s search was very thorough, and had no restrictions on language of publication, study design, or publication status. It also searched for gray literature via published conference proceedings, author correspondence, and bibliographical searches.

Information sources

We included all studies included in Jongsma et al. (2019)'s meta-analysis and all citations listed in the supplemental data provided with Jongsma's study. To cover studies published after Jongsma et al. (2019)'s review, we performed a systematic search for additional studies published up to December 31st 2019.

Search

We used the same search string used by Jongsma et al. (2019), which she adapted from Kirkbride, Errazuriz, et al. (2012), to query PubMed for studies published between January 1st 2018 and December 31st 2019 (see Supplement 4). We performed bibliographic searches whenever possible. We had no language restrictions. We did not query other databases. We did not search the gray literature.

Eligibility criteria

We did not examine studies published before 2008 because one category of interest (applying a cumulative time frame) relies on types of clinical diagnostic practice and electronic data warehouses that only started to emerge at that time. We limited our selection to Northern European studies to reduce potential heterogeneity in health care systems. We considered only incidence estimates for schizophrenia to reduce potential heterogeneity in diagnostic practices.

Therefore, citations were eligible if they contained incidence data, or data from which incidence could be derived (numerator and denominator); included patients (aged 18–64 years) diagnosed with a first episode of schizophrenia; covered populations in Northern Europe; were published between 2008 and 2019, and were listed either in Jongsma et al. (2019)'s meta-analysis (if published 2008–2017) or in PubMed (if published 2018–2019).

Study selection and data collection process

We (AH and SH) first selected on title. We included studies if their title mentioned: (a) 'incidence', 'rate' or 'risk', and (b) one of the words 'schizophrenia', 'psychosis' or 'mental disorders'. We excluded studies with titles referring to specific subgroups as indicated by one of the diagnostic specifiers 'affective', 'postpartum', 'drugs or substance induced psychosis', or subpopulation specifiers 'in or among' 'migrants', 'youth', 'veterans', 'military', 'type 1 diabetes', 'adoptees', 'epilepsy' or 'immune-mediated inflammatory disease'.

We (SH) then selected on the full text. We included studies if they reported estimates of the incidence of 'narrow schizophrenia', defined as 'DSM-IV 295.x' or 'ICD-10 code F20 (including F21 and F25 if possible)' in the general population. We excluded non-European and South European studies to reduce heterogeneity from different healthcare systems and cultural effects on seeking healthcare.

If two or more studies reported on the incidence of schizophrenia in the same population, we included only one. To decide which one, we (SH and AH) assigned priority according to study period (more recent, larger) and quality (more detailed information, state-of-the-art procedures) to arrive at consensus. If two or more methods had been used in the same population, we included one estimate for each method.

Data items

For each study and (if necessary) for each type of study design applied in that study, we collected data related to publication, study period, study population (i.e. country, area, urbanicity, sex and age), study design (i.e. coverage of services, time frame of diagnosis, reliability of diagnosis), and the incidence estimate (i.e. cases and person years at risk).

Coverage of services could range from: (1) 'specialized services' such as Early Psychosis Intervention (EPI) services, and emergency or in-patient services, to the broader set of (2) 'general' psychiatric or addiction services, and further to (3) primary or somatic medical care, and ultimately to (4) the general population. The time frame of diagnosis is the interval between the first contact with a service and the moment a diagnosis is made. It could range from: (1) case ascertainment at first contact, to (2) later stages of treatment, e.g. subjects presenting initially with another diagnosis, ultimately extending to (3) life-time follow-up. Finally, the reliability of diagnosis could range from diagnosis based on: (1) research diagnostic procedures, to (2) clinical criteria diagnoses (e.g. DSM-5 or ICD-10) and (3) non-standardized diagnostic procedures.

Age was categorized according to Howard et al. (2000) in 'early onset' (age < 40 years), 'late onset (age 40–59 years) and 'late onset' (age > 60 years). Urbanicity was classified in three categories: urban, rural, and mixed (i.e. for entire population estimates, such as studies from Denmark). We used the level of urbanicity that each study had assigned to itself.

Assessment of study quality

Kirkbride, Errazuriz, et al. (2012) and Jongsma et al. (2019) used a 7-point quality score. That score was not applicable to our review on 3 out of the 7 points because they relate to the first contact design in particular ('standardized research diagnosis' and 'leakage study') or to studies of risk factors such as ethnicity ('blinding to demographic variables'). For inclusion in our meta-analyses, we required that all studies meet at least all four remaining criteria ('defined catchment area', 'accurate denominator', 'population based case-finding', and 'inclusion criteria'). We nevertheless scored studies on all 7-points for consistency with Kirkbride, Errazuriz, et al. (2012) and Jongsma et al. (2019). For our purposes we considered any study meeting the four core criteria listed above as 'high quality'.

Summary measures

The principal summary measure was the treated incidence rate of schizophrenia per 100 000 person years in the general population.

Synthesis of results

All incidence rates are expressed as number of cases per 100 000 person years. We calculated exact confidence intervals for Poisson rates using the `pois.exact()` function from the 'epitools' package (Aragon et al., 2017) in R version 3.6.1 (R Core Team, 2020).

We calculated pooled incidence rates for each category of study population (i.e. age, urbanicity) and study design (i.e. coverage, time frame, reliability).

We calculated the proportion of between-study variance explained by the covariates by comparing the estimated between-study variance τ^2 , with its value when no covariates are fit τ_0^2 . Adjusted R^2 is the relative reduction in the between study variance $R^2 = \tau_0^2 - \tau^2$ (Harbord & Higgins, 2008).

Additional analyses + meta-regression + sensitivity analysis

To examine how our three design factors related to the incidence, adjusting for differences in population characteristics, we first calculated unadjusted pooled incidence ratios for each of the three variables of interest (coverage, time frame and reliability), and the two covariates (urbanicity and age). Next, to adjust for interdependencies between variables, we conducted a multivariable meta-regression analysis to estimate incidence ratios for

each factor in a single model. To allow for variation both within and between studies, we used a mixed-effects model with restricted maximum likelihood (REML) estimators. We used the Knapp-Hartung adjustment to obtain more reliable confidence intervals (Knapp & Hartung, 2003) and permutation tests to assess the robustness of our model (Higgins & Thompson, 2004). The regression was performed using the ‘meta’ (Balduzzi et al., 2019) and ‘metaphor’ packages (Viechtbauer, 2010) in R.

To rule out bias from including estimates from our own research group (i.e. tilting the scale towards results that confirm our prior findings) we repeated the meta-regression analyses without our own data.

Results

Study selection

The results of the study selection are summarized in a flowchart (Figure 7.1).

Jongsma et al. (2019) identified a total of 125 unique publications between 2008–2017, listed in her review or in the supplement. The search in the PubMed database yielded 527 publications between 2018–2019.

Based on title, we included 70/527 publications from our Pubmed search (left-hand column in the flowchart) and 68/125 publications from Jongsma et al. (2019)’s study (right-hand column) that explicitly mentioned: (a) ‘incidence’, ‘rate’ or ‘risk’, and (b) one of the words ‘schizophrenia’, ‘psychosis’ or ‘mental disorders’. We then excluded 50/70 and 5/68 studies because the titles included the words ‘review’ or ‘meta-analysis’, resulting in respectively 20 and 63 studies. We then excluded 14/20 and 14/63 studies, with titles referring to specific subgroups as indicated by one of the diagnostic specifiers or subpopulation specifiers, resulting in respectively 6 and 49 remaining studies.

Based on the full text, we excluded 4/6 and 9/49 studies from non-European or South-European populations resulting in respectively 2 and 40 remaining studies. We then excluded 1/2 and 18/40 studies because they did not report estimates of the incidence of ‘narrow schizophrenia’. Finally, we excluded 1/1 and 10/22 studies for miscellaneous reasons: two studies that did not describe a general population, one study where coverage and time frame could not be assessed, one study that was a conference abstract, one study with a small population sample ($n < 4000$), and 6 studies that reported duplicate or overlapping findings.

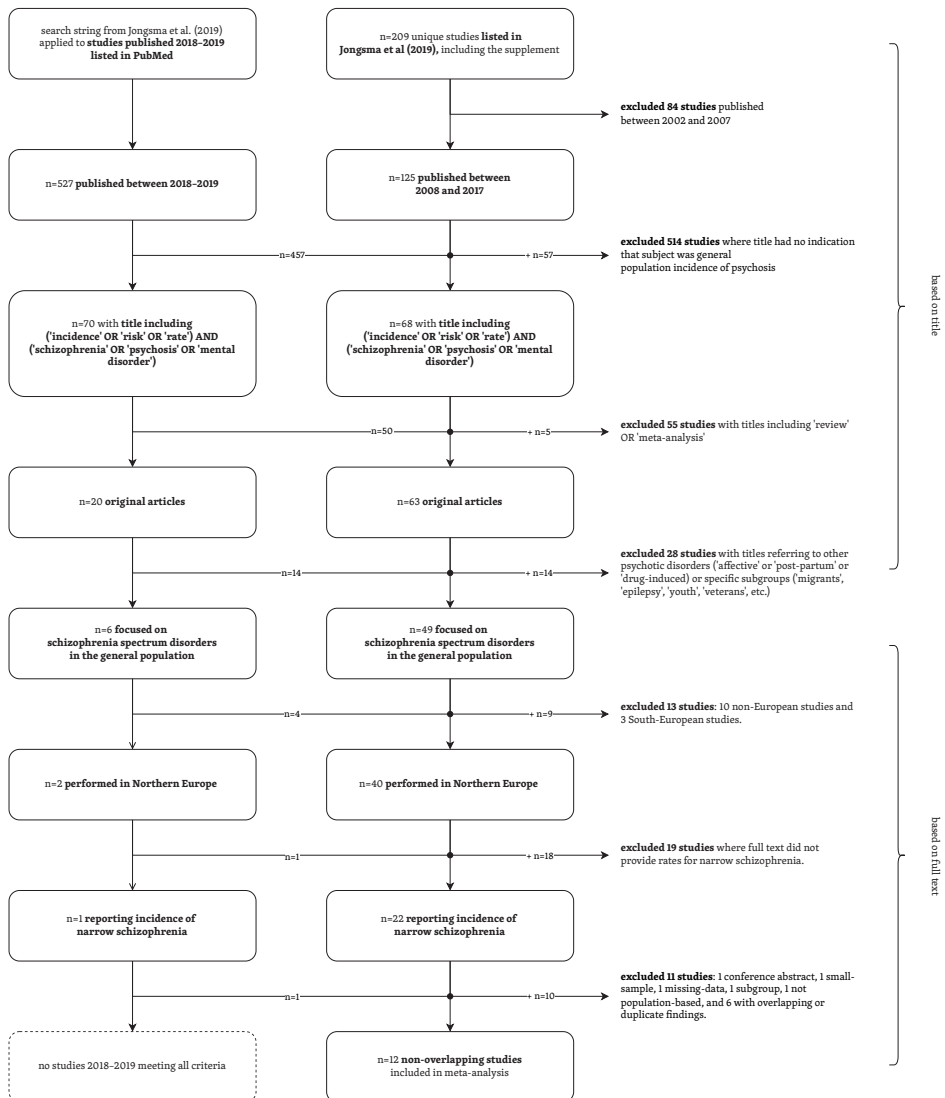


Figure 7.1 PRISMA flowchart

To identify these six duplicates, we determined that the remaining 18 publications described estimates in 15 study populations in four countries, i.e. six from Denmark (Castagnini & Foldager, 2013; Kühl et al., 2016; Nielsen et al., 2017; Paksarian et al., 2015; Sørensen et al., 2015; Vassos et al., 2016), two from Sweden (Jørgensen et al., 2010; Söderlund et al., 2015), two from the Netherlands (Boonstra, Wunderink, Wit, et al., 2008; Hoogerzeil et al., 2014) and five from the United Kingdom (Bhavsar et al., 2014; Kirkbride et al., 2017; Kirkbride et al., 2008; Kirkbride, Jones, et al., 2012; Reay et al., 2010). Although some publications described the same population, the study designs were different and

were therefore included separately in our analysis. All others were treated as duplicates and excluded.

Our selection procedure resulted in a set of 12 publications (Bhavsar et al., 2014; Boonstra, Wunderink, Wit, et al., 2008; Castagnini & Foldager, 2013; Hogerzeil et al., 2014; Jørgensen et al., 2010; Kirkbride et al., 2017; Kirkbride, Jones, et al., 2012; Paksarian et al., 2015; Reay et al., 2010; Salokangas et al., 2010; Sørensen et al., 2015; Szoke et al., 2016). All studies had previously been included in Jongsma et al. (2019)'s meta-analysis. The search for new studies published 2018–2019 identified no new publications meeting all criteria for inclusion in this meta-analysis.

Study characteristics

Table 7.1 shows the incidence estimates with the associated design and population factors. Between 2008 and 2019, 12 European studies together reported 17 estimates of the incidence of schizophrenia in the general population. These studies were from the Netherlands ($n = 2$) (Boonstra, Wunderink, Wit, et al., 2008; Hogerzeil et al., 2014), UK ($n = 4$) (Bhavsar et al., 2014; Kirkbride et al., 2017; Kirkbride, Jones, et al., 2012; Reay et al., 2010), Sweden ($n = 1$) (Jørgensen et al., 2010), Denmark ($n = 3$) (Castagnini & Foldager, 2013; Paksarian et al., 2015; Sørensen et al., 2015), Finland ($n = 1$) (Salokangas et al., 2010) and France ($n = 1$) (Szoke et al., 2016).

All 12 studies were population-based, had specific inclusion criteria, and had an accurate denominator for a defined catchment area, i.e. had a quality score of 4 or higher in terms of Kirkbride, Jones, et al. (2012) 7-point score and were considered 'high quality' for our purposes. Our scores diverged from those by Jongsma et al. (2019) for three studies (Bhavsar et al., 2014; Boonstra, Wunderink, Wit, et al., 2008; Reay et al., 2010) because we classified them as population-based, and as having an accurate denominator. In our sample, seven studies scored 4/7 points, four scored 5/7 points (Bhavsar et al., 2014; Kirkbride et al., 2017; Reay et al., 2010; Szoke et al., 2016), and one (Kirkbride, Jones, et al., 2012) scored 6/7 points. The quality factor 'research diagnosis'—by definition—was always present in our category 'using research diagnosis' and vice-versa. Otherwise, there was no association between study quality and study design, or between study quality and estimate size.

Table 7.1 Characteristics of 12 Northern European studies reporting 17 incidence rates (IR) of schizophrenia, published between 2008 and 2017

	Study quality	Country	Obs	Person years	IR	95% CI	Design factors		Population factors		
							Coverage	Time frame	Diagnosis	Age of onset (yrs)	Density
1	Reay et al. (2010)	UK	60	1 363 485	4.4	3.4 to 5.7	Specialized	First contact	Clinical	< 60	Rural
2	Boonstra et al. (2008)	NL	24	348 222	7.0	4.4 to 10.3	General	First contact	Clinical	< 40	Rural
3	Castagnini and Foldager (2013)	DK	4576	49 921 662	92	8.9 to 9.4	General	First contact	Clinical	< 60	Mixed
4	Szoke et al. (2016)	FR	66	536 168	12.3	9.5 to 15.7	General	First contact	Clinical	< 60	Urban
5a	Jorgensen et al. (2010)	SE	46	348 351	132	9.7 to 17.6	Specialized	First contact	Clinical	< 40	Urban
6	Kirkbride et al. (2017)	UK	350	2 021 663	17.3	15.5 to 19.2	General	First contact	Research	< 40	Mixed
7a	Hogerzeil et al. (2014)	NL	254	1 221 486	20.8	18.3 to 23.5	General	First contact	Research	< 60	Urban
5b	Jorgensen et al. (2010)	SE	96	348 351	27.6	22.3 to 33.7	General	First contact	Clinical	< 40	Urban
8	Sorensen et al. (2015)	DK	15 074	49 898 592	302	29.7 to 30.7	General	Cumulative	Clinical	< 60	Mixed
7b	Hogerzeil et al. (2014)	NL	238	748 423	31.8	27.9 to 36.1	General	First contact	Research	< 40	Urban
9	Kirkbride et al. (2014)	UK	268	828 546	32.3	28.6 to 36.5	General	First contact	Research	< 60	Urban
10a	Salokangas et al. (2011)	Ft	9442	27 661 925	34.1	33.4 to 34.8	Specialized	Cumulative	Clinical	< 60	Mixed
11	Paksarian et al. (2015)	DK	6469	15 800 000	40.9	39.9 to 41.9	General	Cumulative	Clinical	< 40	Mixed
10b	Salokangas et al. (2011)	FI	6988	16 638 863	42.0	41 to 43	Specialized	Cumulative	Clinical	< 40	Mixed
12	Bhaysar et al. (2014)	UK	405	741 758	54.6	49.4 to 60.2	General	First contact	Research	< 40	Urban
7c	Hogerzeil et al. (2014)	NL	843	1 221 486	69.0	64.4 to 73.8	General	Cumulative	Clinical	< 60	Urban
7d	Hogerzeil et al. (2014)	NL	539	748 423	72.0	66.1 to 78.4	General	Cumulative	Clinical	< 40	Urban

Rates were classified according to study design (coverage, time-frame and diagnosis) and population factors (age of onset and urbanicity), and are presented in ascending order of incidence. 'Obs' number of cases, 'IR' incidence rate of narrow schizophrenia.

Estimate characteristics

In total, study selection and data extraction resulted in 17 estimates of the treated incidence of 'narrow schizophrenia' in the general population, for a variety of study designs (i.e. coverage in two levels, time frame in two levels, and reliability in two levels) applies to a variety of study populations (i.e. age in two levels, urbanicity in three levels), adding up to 85 million person years at risk.

This sample contained no estimates in primary care, somatic care, or in the general population. We dropped gender as category for analysis because this information was typically not provided. Information on population characteristics was available for two categories only: age range and urbanicity. There was insufficient information to separate 'early onset' from 'late onset' (40–59 years) and no data were available for 'very late onset' (> 60 years). We therefore merged the age categories into 'early onset' (age < 40 years) and 'early to late onset' (< 60 years). Urbanicity could be assessed for all studies. Three studies (Hogerzeil et al., 2014; Jørgensen et al., 2010; Salokangas et al., 2010) reported estimates based on more than one design or subpopulation and therefore contributed more than one estimate to our data set.

Meta-analysis

Incidence estimates ranged from 4.4 (Reay et al., 2010) to 72.0 (Hogerzeil et al., 2014) per 100 000 person years (median 30; interquartile range 13–41 per 100 000 person years).

The pooled estimate was 40.2 per 100 000 person years (95% confidence interval 39.5 to 40.8) for early onsets (< 40 years) and 23.1 per 100 000 person years (95% confidence interval 22.8 to 23.3) for early-to-late onsets (< 60 years).

Between study heterogeneity (I^2) in the study sample was 98.7% and 99.9% for early and early-to-late onsets, respectively.

Meta-regression

Unadjusted pooled incidences and incidence ratios for individual factors in study design or study population are shown in Table 7.2. In this single variable comparison, no significant differences were found for coverage of services, quality of diagnosis, or age of onset. For 'time frame for diagnosis', the incidence estimates were more than threefold higher for cumulative time frames versus first contact studies (incidence ratio 3.21; 95% CI 3.13 to 3.30). In addition, estimates from rural populations were roughly six-fold lower than in urban populations (0.12; 95% CI 0.10–0.15).

Table 7.2 Unadjusted pooled incidence rates and incidence rate ratios for individual aspects of study design or type of study population

	N	Obs	Person years	IR	95% CI	unadj IRR	95% CI
Coverage of services	4	16 536	4 601 2624	35.9	35.4 to 36.5	ref	
Specialized services	13	29 202	12 4384 780	23.5	23.2 to 23.7	0.65	0.64 to 0.67
General psychiatric services	11	6 383	58 428 115	10.9	10.7 to 11.2	ref	
Time frame to diagnosis	6	39 355	111 969 289	35.1	34.8 to 35.5	3.21	3.13 to 3.30
First contact	5	1 515	5 561 876	27.2	25.9 to 28.6	ref	
Cumulative	12	44 223	164 835 528	26.8	26.6 to 27.1	0.98	0.94 to 1.04
Quality of diagnosis	8	30 583	132 653 350	23.1	22.8 to 23.3	ref	
Research diagnosis	9	15 155	37 744 054	40.2	39.5 to 40.8	1.74	1.71 to 1.78
Clinical diagnosis	9	2 755	6 742 992	40.9	39.3 to 42.4	ref	
Age of onset	6	42 899	161 942 705	26.5	26.2 to 26.7	0.65	0.62 to 0.67
Early-to-late onset	2	84	1 711 707	4.9	3.9 to 6.1	0.12	0.10 to 0.15
Early onset							
Urbanicity							
Urban							
Mixed							
Rural							

'N' number of estimates, 'Obs' number of observed cases, 'IR' incidence rate per 100 000 person years, 'IRR' incidence exact Poisson confidence interval

Results of our multivariable meta-regression analysis are presented in Table 7.3. The meta-regression indicated that among adults aged 15–59 years in a general urban population, a study using research diagnoses made in specialized services and applying a first contact time frame would estimate the incidence of schizophrenia at 25 per 100 000 person years (Knapp-Hartung adjusted 95% CI 15 to 40). But in the same population—a study using clinical diagnoses would report a 0.55 (0.38 to 0.81) times lower estimate, and one applying a cumulative time frame would report a 4.04 (2.78 to 5.87) times higher estimate. If the same design were used in mixed and rural settings, estimates would be 0.54 (0.39 to 0.75) and 0.33 (0.18 to 0.6) times lower, respectively. If age of onset were to be restricted to early age of onset, the estimate would be 1.34 (1.02 to 1.75) times higher. Extending coverage to general psychiatric services would not increase estimates significantly (1.12 times; 95% CI 0.88 to 1.43).

The three study design factors together explained 67% of between study variance (adjusted R^2). A complete model, including the two differences in study population explained 91% of between-study variance.

Permutation tests confirmed that the estimators were robust. Running the meta-regression on subsets (i.e. the set of estimates reporting ‘early onset’ and the set of estimates for ‘early-to-late onset’ separately) did not change the outcome. Likewise, removing our own data (Hogerzeil et al., 2014) did not change the outcome.

Table 7.3 Multivariate meta-regression analysis, modeling incidence rate ratios' estimates in terms of coverage of service, time frame for diagnosis, and quality of diagnosis, with urbanicity and age as covariates

		IRR	95% CI (unadjusted)	p.crude	95% CI adj1	p.adj1	p.adj2
Coverage of services	Specialized services	1.00	ref		ref		
	General psychiatric services	1.12	(0.88 to 1.43)	0.354	(0.78 to 1.61)	0.495	0.501
Time frame to diagnosis	First contact	1.00	ref		ref		
	Cumulative	4.04	(3.14 to 5.2)	0.000	(2.78 to 5.87)	0.000	0.001
Quality of diagnosis	Research diagnosis	1.00	ref		ref		
	Clinical diagnosis	0.55	(0.43 to 0.72)	0.000	(0.38 to 0.81)	0.007	0.008
Age of onset	Early-to-late onset	1.00	ref		ref		
	Early onset	1.34	(1.12 to 1.6)	0.002	(1.02 to 1.75)	0.036	0.036
Urbanicity	Urban	1.00	ref		ref		
	Mixed	0.54	(0.44 to 0.67)	0.000	(0.39 to 0.74)	0.001	0.001
	Rural	0.33	(0.22 to 0.49)	0.000	(0.18 to 0.6)	0.002	0.003

The intercept for the meta-regression model was 25 per 100 000 person years (95% CI 15 to 40). 'IR' incidence rate per 100 000 person years, 'IRR' incidence rate ratio, 'ref' reference category, '95% CI adj1' 95% confidence interval after Knapp-Hartung adjustment, 'p.adj1' p value after Knapp-Hartung adjustment, 'p.adj2' p value after permutation test.

Discussion

We conducted a review of 12 selected studies on the incidence of narrow schizophrenia in the general adult population published between January 1st 2008 and December 31st 2019. We examined the impact of differences in study design on the variation of reported incidences. We found 17 estimates in six countries, covering more than 15 000 individual cases and 85 million person years.

We examined the impact of three study design characteristics (coverage, time frame, reliability of diagnosis), adjusting for population characteristics with two covariates (age, urbanicity). Differences in study design together explained 67% of between-study variation, while a more complete model, including age and urbanicity as covariates, explained 91%. In our model, a longer 'time frame' resulted in four-fold higher estimates, and clinical diagnoses, compared to standardized research diagnoses, reduced estimates by half.

The four-fold difference between estimates based on cumulative vs. first contact time frames is in line with our previous study, where we compared a cumulative case-register design to a first contact design in a single population (in the Netherlands), which demonstrated a 3.3-fold higher estimate for the cumulative time frame (Hogerzeil et al., 2014). Similarly, other case-register studies have tended to report higher incidence estimates than first contact studies (Anderson et al., 2018; Jongsma et al., 2018; Kirkbride, Errazuriz, et al., 2012; McGrath et al., 2004; Pedersen et al., 2014; Thorup et al., 2007). The findings in this study agree with our previous findings (Hogerzeil et al., 2014) and confirm them independently since our conclusions did not change when we removed our own data from the analysis. They confirm the threefold difference between register studies and first contact studies reported in Jongsma et al. (2019)'s meta-analysis. They expand on her finding by untangling the relative contributions of separate design aspects.

One limitation of the first contact approach as commonly practiced is that it cannot account for long delays in reaching an ultimate diagnosis of schizophrenia. Most patients with schizophrenia first report to services with other symptoms, such as depression, anxiety or substance abuse (Hogerzeil et al., 2014; Rietdijk et al., 2011; Simon et al., 2017). They may also present with psychotic symptoms, but not per se schizophrenia. In our prior study (Hogerzeil et al., 2014), the median interval between first contact and the index diagnosis of schizophrenia was 4.9 years (interquartile range 1.1 to 8.8), but the interval sometimes extended beyond 25 years. In theory, Jablensky et al. (1992)'s original first contact inclusion criteria do not exclude patients who first contacted services for other reasons. But in practice, most first contact studies have not actively screened for

onsets of schizophrenia among patients contacting services for other reasons, or patients currently under treatment for other reasons than psychosis.

A criticism on our approach could be that we focus our review on narrowly-defined schizophrenia. Many first contact studies nowadays are performed in Early Intervention services, as close as possible to the emergence of psychotic symptoms. Such services tend to work with provisional clinical diagnoses such as ‘psychosis NOS’ (not reviewed here), of which many are perhaps ultimately diagnosed with schizophrenia at later stages of treatment. So they treat more (future) cases of schizophrenia than is reflected in their provisional numbers. Our focus on narrow-schizophrenia therefore favors case registers compared to first contact studies because registers work with ultimate rather than provisional diagnoses. Although the criticism can be a valid explanation for lower incidence estimates in first contact studies, it also underscores the potential under detection of true cases of narrow-schizophrenia in such designs.

Prior work suggests that both the primary care system and general psychiatric services play an important role in first diagnosis of psychotic disorder, and these physicians may be involved in ongoing psychiatric care, especially in settings where specialized services are unavailable (Anderson et al., 2018; Rietdijk et al., 2011; Simon et al., 2017). We had no data on the incidence of schizophrenia in primary care, somatic medical care or the general population. But contrary to our expectation, we found no differences between specialized vs general psychiatric services as channels for case-detection. This has implications for healthcare: in that increasing service coverage (beyond services typically used by psychotic patients) to detect more cases of incident schizophrenia will not result in better estimates, if the time frame remains limited to diagnosis at first contact. One explanation could be that every subject with clinically relevant schizophrenia is eventually referred to specialized services (Weiser et al., 2012), and can be counted at that later point in time if the study design allows for such a pathway to care.

The two-fold difference between estimates based on research diagnoses vs. clinical diagnoses was also unexpected. It runs counter to common intuition that clinicians diagnose schizophrenia too easily and that relying on (presumably) conservative, standardized research procedures would result in lower (but more valid) incidence estimates (Castillejos et al., 2019; Jongsma et al., 2019). The idea that research diagnoses are to be preferred over clinical diagnoses is contradicted by reports that clinical diagnoses can be valid (Dalman et al., 2002; Ekholm et al., 2005; Ludvigsson et al., 2011; Uggerby et al., 2013) and stable over time (Fusar-Poli et al., 2016). Because our sample contained no ePCR studies with research diagnoses, comparisons between studies based on clinical vs. research diagnoses were restricted to first contact studies. The counter-intuitive

finding therefore bears primarily on first contact studies. It offers a new perspective, by suggesting that clinicians may in fact be more conservative than researchers in diagnosing schizophrenia. We speculate that clinicians are reluctant to diagnose schizophrenia formally to avoid the stigma associated with the label.

Limitations

The large attrition of eligible studies was a consequence of the quality criteria adopted to answer our research question. We restricted our search to studies published from 2008 onwards because clinical practices have become more standardized and electronic patients records better available in recent years. The further restriction to studies from Northern Europe resulted in a high-quality study sample that was comparable in terms of culture and health systems. Despite the small number of studies, the sample still covered 85 million person years and more than 15 000 cases of schizophrenia.

The risk of bias is lower for incidence studies than for RCTs. They are not blinded or randomized. There are no financial or ideological incentives to distort the incidence estimate, or the association between study method and incidence. The quality scores of the studies included in our review were high and not related to estimate size. Our update for the years 2018–2019 did not include the gray literature, however, and we did not query databases other than Pubmed. But arguably studies not listed on Pubmed are no different with respect to our main finding.

Another limitation is that our information on population characteristics was only for age and urbanicity. We had limited information on relevant age bands and no information on gender, ethnicity or other socio-economic or biological risk factors. Despite this limitation, including age and urbanicity as covariates in our final regression model explained 91% of between-study variation. This may be due to the homogenous selection of studies (all from Northern Europe), which was helpful to demonstrate the specific contribution of design factors. But our findings underestimate the contribution of other population characteristics as a source of variation.

Finally, in our statistical model, we did not account for interactions between study design characteristics and population characteristics. Such interactions are plausible, e.g. older, non-migrant females with mild symptoms are less likely to be included in first contact studies than young migrant males with acute onset of psychosis (Hogerzeil et al., 2016). In our analyses, the net effect would be conservative, e.g. selection bias in first contact studies in favor of including subjects with higher incidence of schizophrenia would shrink the contrast with ePCRs observed in this study.

Conclusions

In conclusion, our selective review demonstrates that differences in study design explain most of the wide variation in reported estimates of the incidence of schizophrenia. This artefact can eclipse true but smaller variations in population risk factors such as gender, age and migrant status. To distinguish cause from noise, future systematic reviews should apply standardized categorizations by type of design (Edwards et al., 2019; Hogerzeil & Hemert, 2019).

