# One-year effect of changing duration of untreated psychosis in a single catchment area

TOR K. LARSEN, INGRID MELLE, SVEIN FRIIS, INGE JOA, JAN OLAV JOHANNESSEN, STEIN OPJORDSMOEN, ERIK SIMONSEN, PER VAGLUM and THOMAS H. McGLASHAN

**Background**    There is highly replicated positive correlation between longer duration of untreated psychosis and poorer outcome.

**Aims**    To study the effect of early intervention in first psychosis on one-year outcome using an historical quasi-experimental design.

**Method**    We compare the outcome of two samples of first-episode psychosis from the same healthcare district at different time periods. The historical control sample was assessed during 1993–1994, before the establishment of a system for early detection of psychosis. The experimental sample is the early detection sample in the Early Treatment and Intervention in Psychosis study assessed during 1997–2000.

**Results**    At 1-year follow-up, the early detection group was younger, had a smaller fraction of individuals with schizophrenia, had less severe negative and general symptoms and had more friends in the past year than the historical control group. No differences were found in clinical course (remission, relapse, continuously psychotic) or positive symptoms, but more patients in the early detection sample were treated as outpatients without hospitalisation.

**Conclusions**    Early detection of schizophrenia in one healthcare sector is associated with less severe deterioration at 1 year.

**Declaration of interest**    None.

Early detection of psychosis has become the focus of much investigation because of the highly replicated positive correlation between longer duration of untreated psychosis (DUP) and poorer outcome (Marshall *et al*, 2005; Perkins, 2006).

The Early Treatment and Intervention in Psychosis study (TIPS) is the first to reduce DUP with an early detection programme involving the creation of easy access psychosis-detection teams and a massive and persistent educational campaign about the first signs of psychosis aimed at the general public, the primary health services and the school system (Melle *et al*, 2004). The study also compared the effect of reducing DUP on the severity of first-episode psychosis using a parallel-control, quasi-experimental design. Such a design compares the clinical presentation and course of patients with first-episode schizophrenia in different healthcare sectors, an experimental sector with an early detection programme (Rogaland, Norway) and two control sectors without such a programme (Ulleval, Norway and Roskilde, Denmark). The term parallel control derives from the fact that the experimental and control sectors collect their samples over the same (parallel) time period; for the TIPS this was 1997–2000 (Melle *et al*, 2004).

In TIPS DUP was reduced significantly in the experimental early detection sector compared to no-early detection control sectors (5 weeks median compared to 16 weeks median). At intake (baseline), patients in the early detection sector had significantly less severe positive, negative and general symptoms ($P < 0.01$ for all comparisons) (Melle *et al*, 2006). At 1-year follow-up the differences in positive and general symptoms disappeared but the negative symptom differences remained significant ($P < 0.005$) (Larsen *et al*, 2006).

The focus of this paper will be a comparison of the TIPS 1997–2000 early detection or experimental sector sample with a different no-early detection control group. This control group consists of all patients with first-episode non-affective psychosis who came to treatment in the same healthcare sector, i.e., Rogaland County, but at an earlier time period, the years 1993–1994, before any educational campaigns. Before the TIPS programme began, this study was carried out to measure DUP in first-episode psychosis in the middle and southern sections of Rogaland County. The investigation used essentially the same inclusion criteria, assessment battery and follow-up procedures as in TIPS (Larsen *et al*, 2001). This sample, coming from the same early detection sector but from a different time before the early detection campaign, provides a no-early detection historical control sample to the 1997–2000 early detection experimental sample.

We have previously reported a comparison between the TIPS early detection sample from the first 2 years of recruitment (1997–1998, $n=66$) with the 1993–1994 historical control sample regarding baseline characteristics ($n=43$) (Larsen *et al*. 2001). We found that the patients with early detection at baseline had a significantly shorter DUP, were younger, misused substances more often, had better premorbid adjustment and had less symptoms. In this paper we compare the 1-year outcome of the full 4-year early detection sample in TIPS (1997–2000, $n=133$) with the 1993–1994 historical control sample at baseline and 1-year outcome. We aim to see if the 1-year outcome findings are similar to those of the TIPS parallel-control study (as reported in Larsen *et al*, 2006) using a different no-early detection historical control group. The null hypothesis is that the findings of the parallel control 1-year outcome will not be replicated.

## METHOD

### Participants

Patients were included during two time periods; 1993–1994 and 1997–2000. The population is 260 000 (historical control period) and 290 000 (early detection period). For both periods the criteria for inclusion were (a) a first episode of a non-affective psychosis according to DSM–IV, i.e. schizophrenia, schizophreniform psychosis, schizoaffective psychosis, delusional disorder, brief psychosis, and psychosis not otherwise specified (NOS) (affective disorder with mood incongruent delusions

was not included in 1993–1994 and is therefore excluded from this report); (b) living in the catchment area; (c) age 15–65 years; (d) IQ >70. The exclusion criteria included having an organic/substance-induced psychosis and/or having received adequate prior antipsychotic treatment. Written informed consent for the follow-up study was obtained from all subjects and the regional ethical research committees approved the study.

## Assessments

The Structured Clinical Interview for DSM–IV (SCID) (Spitzer et al, 1995) was used in 1997–2000. SCID for DSM–III–R was used in the 1993–1994 period; the diagnosis were converted into DSM–IV diagnosis by T.K.L. Premorbid functioning was measured by the Premorbid Adjustment Scale (PAS), which describes four premorbid periods in life: childhood (≤11 years), early adolescence (12–15 years), late adolescence (16–18 years), and adulthood (≥19 years) (Cannon-Spoor et al, 1982). A previous analysis identified two premorbid dimensions (a) social, consisting of PAS items, social isolation and peer relationships; and (b) academic, containing school performance and school adaptation. Two parameters for each dimension are rated: (a) childhood level of adjustment; and (b) degree of change of level of adjustment over post-childhood developmental phases (for details about this modification see Larsen et al, 2004).

The DUP was measured as the time from onset of psychosis until the start of adequate treatment (for details see Melle et al, 2004). Level of symptoms were measured by the Positive and Negative Syndrome Scale (PANSS; Kay et al, 1987). Misuse of alcohol and other drugs was measured by the Drake Scale (Drake et al, 1990). Social functioning was measured with the Strauss–Carpenter Scale (Strauss & Carpenter, Jr, 1974). At 1 year all assessments were repeated, including the SCID.

All major baseline assessments such as diagnosis, PANSS, Global Assessment of Functioning (GAF), drug misuse and DUP underwent tests of intra- and inter-site reliability with satisfactory results within the early detection period (for details see Friis et al, 2003). For the historical control period reliability tests were carried out for diagnosis (20 written case vignettes rated by T.K.L. and S.O.) with a kappa of 0.89.

For the PANSS and GAF a similar study was carried out with intraclass correlations (ICC 1.1) between 0.65 and 0.90. A test–retest study was carried out for PAS and DUP with satisfactory results (for details see Larsen et al, 1996).

## Data analysis

### Statistical procedures

Analyses were performed with the statistical package SPSS (version 12.0 for Windows). The applied methods are reported for all group comparisons. All tests were two-tailed. We used nonparametric tests for data without normal distribution. As noted in several other studies, the DUP does not seem to have a normal distribution, whereas its natural logarithm does. In multiple linear regression analyses DUP was transformed to its natural logarithm.

For the multivariate analysis the selection of variables to be included in the regression model was based on their assumed clinical importance through reviews of relevant studies and a set of variables comprising measures of differences between the areas. The variables were entered hierarchically with 'from the early detection sample' on the last step. The final model was examined for interaction effects and the effects of outliers and influential observations, including leverages.

## RESULTS

### Baseline findings (Table I)

In the historical control sample 43 patients were included and in the early detection period 118. The historical control inclusion period was 2 years, the early detection period twice as long. The number of included patients during the early detection period showed an increase of 37%. At baseline the historical control sample was older and had poorer premorbid social functioning expressed as a greater deterioration in social functioning from childhood to adulthood. This sample had a longer DUP, more individuals with schizophrenia, more severe positive, negative and general symptoms on the PANSS, and drug misuse was less severe.

### Follow-up findings (Table 2)

At 1-year follow-up we reassessed 100% of the historical control sample for all variables, for the early detection sample we were

able to reassess 99% for illness-course, 97% for diagnosis and 88% with PANSS. Fewer of the patients in the early detection group were hospitalised, but the 1-year clinical course was not different between either of the two samples. The historical sample had a higher fraction of individuals with schizophrenia, more negative and general symptoms on the PANSS, and also had fewer friends in the past year.

The multivariate analysis revealed that negative symptoms were related significantly to social deficits in childhood, to increasing social deficits during subsequent premorbid development, and to developing a narrow schizophrenia disorder. Reducing DUP, i.e. coming from the early detection period (1997–2000), remained significant after controlling for the other variables.

## DISCUSSION

The findings of the study are that the TIPS early detection programme in the middle and southern sectors of Rogaland County succeeded in significantly reducing DUP compared to an historical control sample from the same sectors.

At baseline, the early detection sample was larger in number, younger in age and displayed healthier premorbid social functioning. The fraction of this sample meeting criteria for DSM–IV schizophrenia was smaller and symptomatically the sample had less severe positive, negative and general symptoms. The only way the early detection sample may have been 'worse' than the historical control sample was in scoring higher on drug misuse.

By 1-year follow-up significantly more of the patients in the early detection group were treated as out-patients, but the clinical course between early detection and historical control was comparable, with similar sample fractions being in remission, in relapse or continuously psychotic. The historical control sample still had a significantly higher fraction of patients with schizophrenia, but the difference between the two groups was less than at baseline, owing primarily to a larger shift in diagnosis from schizophreniform to schizophrenia within the early detection sample. Symptomatically at 1 year there were no longer differences in positive symptoms, but the differences in negative and general symptoms remained robust, particularly the former. Individuals with early detection at 1 year scored higher for having friends, a

**Table I** Comparison between historical control sample and early detection sample with first-episode psychosis at baseline

| | Historical control group[1] | Early-detection group[1] | P |
|---|---|---|---|
| Gender, % female | 35 | 41 | 0.6 |
| Age, years | 28.4 (8.3) | 24.8 (7.9) | 0.02 |
| Premorbid Adjustment Scale[2], score | | | |
| Social factor | | | |
| Childhood | 1.3 (1.0) | 1.0 (1.3) | 0.1 |
| Degree of change scores | 1.5 (1.3) | 0.9 (1.7) | 0.03 |
| Academic factor | | | |
| Childhood | 2.0 (1.1) | 1.7 (1.1) | 0.8 |
| Degree of change scores | 0.7 (1.2) | 0.7 (1.3) | 0.8 |
| Duration of Untreated Psychosis[3] mean (median) | 114.2 (26.0) | 28.2 (6.0) | 0.0001 |
| Diagnosis[4] (%) | | | |
| Schizophrenia | 77 | 32 | 0.0001[1] |
| Schizophreniform | 7 | 27 | |
| Schizoaffective | 2 | 17 | |
| Delusional disorder | 12 | 5 | |
| Brief psychosis | 2 | 9 | |
| Psychosis not otherwise specified | 0 | 9 | |
| Positive and Negative Syndrome Scale[5], score (s.d.) | | | |
| Positive symptoms | 21.4 (4.8) | 18.7 (5.1) | 0.003 |
| Negative | 19.8 (6.8) | 14.8 (6.7) | 0.0001 |
| General | 40.5 (8.9) | 31.6 (8.4) | 0.0001 |
| Misuse[6], Drake Scale, score (s.d.) | | | |
| Drugs | 1.3 (0.6) | 1.9 (1.1) | 0.001 |
| Alcohol | 1.8 (0.6) | 1.9 (0.7) | 0.2 |
| Strauss–Carpenter[7], mean (s.d.) score | | | |
| Work in the past year | 1.9 (1.7) | 2.0 (1.8) | 0.6 |
| Friends last year | 2.5 (1.5) | 3.0 (1.3) | 0.1 |

$\chi^2$ between historical control group and early detection group regarding schizophrenia or other types of psychosis.
1. $n=43$ for historical control (HC) sample, $n=118$ for early detection (ED) sample.
2. $n=42$ for HC, $n=117$ for early detection (ED) sample.
3. $n=43$ for HC, $n=118$ for ED.
4. $n=43$ for HC, $n=118$ for ED.
5. $n=43$ for HC, $n=117$ for ED.
6. $n=32$ for HC, $n=117$ for ED.
7. $n=43$ for HC, $n=115$ for ED.

new finding given that they were not different on this parameter at baseline.

Overall, psychopathologically the differences between the the two sample groups samples were more attenuated at 1 year than they were at baseline. Although this suggests that with more time the early detection sample might 'catch up' with the historical control sample *vis à vis* level of deterioration and chronicity, many of the initial differences remained robust at 1 year. This was especially true for the difference in negative symptoms between the groups, which was virtually the same at 1 year as it was at baseline. Furthermore, when we analysed the difference for negative symptoms for a subsample with core and narrow schizophrenia, the difference persisted even for those with poor premorbid social adjustment, suggesting that early detection may attenuate primary as well as secondary negative symptoms (Kirkpatrick *et al*, 1989).

The 1-year differences in favour of the early detection sample in this study do not support the null hypothesis. On the other hand we cannot say our findings replicate the parallel-control study's 1-year outcome (Larsen *et al*, 2006) because the experimental samples are not completely different between studies. Nevertheless we can say our early detection–historical control differences

in this historical control study are more numerous and quantitatively more robust than they are in the parallel-control study. This may reflect that the experimental-control differences in DUP in this study versus the parallel-control study are also more robust, e.g. 6 *v.* 26 weeks (median) in this study compared to 5 *v.* 16 weeks (median) in the parallel-control study (Melle *et al*, 2004).

The patients in the historical control group clearly began first treatment for psychosis on average much later in their psychosis than the early detection patients. The patients in this control sample may, in effect, be further along in the window of deterioration that characterises the longitudinal course of schizophrenia (McGlashan, 1988; McGlashan & Fenton, 1993). If so, they may also be closer to the plateau in deficit formation that determines the end of this window. Given sufficient time and more follow-ups, the early detection sample of this study may eventually reach the same level of deficit as the historical control sample if early detection does nothing to reduce the course of deterioration in psychosis. If such is the case, then the differences we are reporting here at 1 year could be an example of a 'lead-time bias' as described by Black & Welch (1993), that is, an artifact of early detection of disorder. If, however, secondary prevention is conferred by early intervention, then the early detection baseline advantage should persist through the window of deterioration and emerge 1–3 years later as a permanent difference or advantage that lasts a lifetime. In both this historical control study and the TIPS parallel-control study, secondary prevention may be happening, especially with negative symptoms, but further follow-up is necessary to rule out the competing hypothesis that what we are seeing is the consequence of lag-times between samples in their progression through a common window of deterioration to an equivalent deficit plateau.

Eearly detection in this study may or may not confer secondary prevention in psychosis, but it clearly provides several examples of tertiary prevention. The early detection group was younger, better functioning premorbidly, and less symptomatic, with fewer individuals meeting criteria for 'core' schizophrenia. If early detection works by reducing the threshold at which the signs and symptoms of psychosis are recognised, then we would expect to see a

**Table 2** Comparison between historical control and early detection samples with first episode psychosis at 1 year follow-up

| | Historical control group | Eearly-detection group | P |
|---|---|---|---|
| Hospitalised[1,2], % | 98 | 76 | 0.001 |
| Course during 1-year follow-up[1,3] (% yes) | | | |
|    In remission (> 1 week) | 56 | 62 | |
|    In relapse | 19 | 10 | |
|    Continously psychotic | 26 | 28 | 0.8 |
| Diagnosis[1,4], % | | | |
|    Schizophrenia | 81 | 54 | 0.001[1] |
|    Schizophreniform | 2 | 6 | |
|    Schizoaffective | 2 | 18 | |
|    Affective mood incongruent | 0 | 5 | |
|    Delusional disorder | 9 | 2 | |
|    Brief psychosis | 4 | 7 | |
|    Psychosis not otherwise specified | 0 | 8 | |
| Positive and Negative Syndrome Scale[1,5], score (s.d.) | | | |
|    Positive symptoms | 12.5 (4.6) | 12.4 (5.6) | 0.9 |
|    Negative | 17.0 (6.8) | 12.2 (5.6) | 0.0001 |
|    General | 29.4 (9.1) | 25.0 (8.9) | 0.008 |
| Strauss-Carpenter[1,6], mean score (s.d.) | | | |
|    Work in the past year | 1.3 (1.5) | 1.3 (1.6) | 1.0 |
|    Friends in the past year | 2.4 (1.5) | 3.0 (1.4) | 0.02 |
|    Symptoms in the past month | 2.4 (0.8) | 2.2 (1.3) | 0.4 |

$\chi^2$ between historical control group and early detection group regarding schizophrenia or other types of psychosis.
1. $n=43$ for historical group.
2. $n=118$ for early detection group.
3. $n=117$ for early detection group.
4. $n=115$ for early detection group.
5. $n=104$ for early detection group.
6. $n=109$ for early detection group.

**Table 3** Multiple linear regression analysis of the effect of independent variables on negative symptoms at 1-year follow-up[1]

| Variables | Coefficients | Confidence intervals | P |
|---|---|---|---|
| (Constant) | | −0.2; 7.4 | 0.64 |
| Age in years | −0.1 | −0.1; 0.1 | 0.86 |
| Gender | −0.004 | −1.9; 1.8 | 0.96 |
| Premorbid social functioning childhood | 0.3 | 0.5; 2.1 | 0.001 |
| Premorbid social functioning change score | 0.4 | 1.0; 2.2 | 0.0001 |
| Log transformed | 0.1 | −0.2; 1.0 | 0.178 |
| Narrow schizophrenia | 0.2 | 1.0; 5.3 | 0.005 |
| Coming from early detection period | 0.2 | 0.7; 4.7 | 0.009 |

1. Model summary: $R^2=0.403$; Adjusted $R^2=0.372$.

greater 'capture' of younger and less severe cases who escape detection until later when they are more obviously psychopathologic. Indeed, one consequence of effective early detection may be to appear to increase the incidence of schizophrenia in a sector when, in effect, all cases are legitimate but

a higher fraction of all possible cases are being detected.

Almost 1 in 4 of the patients in the early detection group was treated for first psychosis as out-patients. Another positive result of early detection may be identifying people as psychotic at a milder stage of

disorder and avoiding hospitalisation, i.e. reducing the 'collateral damage' and expense that usually attends a first psychotic episode.

## Strengths of the study

A strength of the study is the well-organised, nationalised mental healthcare system of Rogaland County, Norway. The samples are representative and are drawn consecutively from the same catchment area separated by only 4 years. No other services treat patients with first-episode psychosis within this region, and the samples have a high follow-up rate.

## Limitations

A major limitation of the study is the (necessary) quasi-experimental design and the probability of cohort effects, i.e. sample differences emerging because of changes over time in the populations studied. For example, our early detection sample had more substance misuse than the historical control sample. We deduce that this difference is real and reflects an epidemic of drug use that arose in Scandinavia around the time of the TIPS project. This finding illustrates how rapidly cohort effects (e.g. differences) can happen, even in cohorts from the same catchment area separated by only 4 years.

Another limitation is that some of the designs and methods were not identical across the early detection and historical control cohorts. For example, time to remission was not rated in the latter cohort. Also, treatment was provided to the patients in the historical control sample during the year but it was not standardized or recorded, so we cannot estimate how much the differences between the samples at 1 year derive from treatment differences. All of the clinical ratings in the historical control cohort were performed by one person (T.K.L.) and, while he was adequately trained, standard interrater reliability between the historical control and the early detection sample was never established. The detection of cases was also more intensive during the TIPS period; in particular, cases from outpatients' clinics might have been overlooked in the historical control sample. Finally, the historical control cohort was assessed diagnostically prior to DSM–IV, consequently, requiring that the DSM–III–R diagnoses be retrospectively reassessed as to whether or not they met DSM–IV criteria. This was done by one

person (T.K.L.) without subjecting the process to reliability testing or consensus determination.

## AKNOWLEDGEMENTS

## REFERENCES

Black, W. C. & Welch, H. G. (1993) **Advances in diagnostic imaging and overestimations of disease prevalence and the benefits of therapy.** New England Journal of Medicine, 328, **1237–1243.**

Cannon-Spoor, H. E., Potkin, S. G. & Wyatt, R. J. (1982) Measurement of premorbid adjustment in chronic schizophrenia. *Schizophrenia Bulletin*, **8**, 470–484.

Drake, R. E., Osher, F. C., Noordsy, D. L., *et al* (1990) Diagnosis of alcohol use disorders in schizophrenia. *Schizophrenia Bulletin*, **16**, 57–67.

Friis, S., Larsen, T. K., Melle, I., *et al* (2003) Methodological pitfalls in early detection studies – the NAPE Lecture 2002. Nordic Association for Psychiatric Epidemiology. *Acta Psychiatrica Scandinavica*, **107**, 3–9.

Kay, S. R., Fiszbein, A. & Opler, L. A. (1987) The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia Bulletin*, **13**, 261–276.

Kirkpatrick, B., Buchanan, R. W., McKenney, P. D., *et al* (1989) The Schedule for the Deficit syndrome: an instrument for research in schizophrenia. *Psychiatry Research*, **30**, 119–123.

Larsen, T. K., McGlashan, T. H. & Moe, L. C. (1996) First-episode schizophrenia: I. Early course parameters. *Schizophrenia Bulletin*, **22**, 241–256.

Larsen, T. K., McGlashan, T. H., Johannessen, J. O., *et al* (2001) Shortened duration of untreated first episode of psychosis: changes in patient characteristics at treatment. *American Journal of Psychiatry*, **158**, 1917–1919.

Larsen, T. K., Friis, S., Haahr, U., *et al* (2004) Premorbid adjustment in first-episode non-affective psychosis: distinct patterns of pre-onset course. *British Journal of Psychiatry*, **185**, 108–115.

Larsen, T. K., Melle, I., Auestad, B., *et al* (2006) Early detection of first episode psychosis; the effect on 1-year outcome. *Schizophrenia Bulletin*, **32**, 758–764.

Marshall, M., Lewis, S. W., Lockwood, A., *et al* (2005) Association between duration of untreated psychosis and outcome in cohorts of first episode patients: a systematic review. *Archives of General Psychiatry*, **62**, 975–983.

McGlashan, T. H. (1988) A selective review of recent North American long-term followup studies of schizophrenia. *Schizophrenia Bulletin*, **14**, 515–542.

McGlashan, T. H. & Fenton, W. S. (1993) Subtype progression and pathophysiologic deterioration in early schizophrenia. *Schizophrenia Bulletin*, **19**, 71–84.

Melle, I., Larsen, T. K., Haahr, U., *et al* (2004) Reducing the duration of untreated psychosis in first episode psychosis: effects on clinical presentation. *Archives of General Psychiatry*, **61**, 143–150.

Melle, I., Johannesen, J. O., Friis, S., *et al* (2006) Early detection of the first episode of schizophrenia and suicidal behavior. *American Journal of Psychiatry*, **163**, 800–804.

Perkins, D. O. (2006) Review: longer duration of untreated psychosis is associated with worse outcome in people with first episode psychosis. *Evidence Based Mental Health*, **9**, 36.

Spitzer, R. L., Williams, J. B. W., Gibbon, M., *et al* (1995) *Structured Clinical Interview for DSM–IV (SCID).* New York State Psychiatric Institute, Biometrics Research.

Strauss, J. S. & Carpenter, W. T., Jr (1974) The prediction of outcome in schizophrenia. II. Relationships between predictor and outcome variables: a report from the WHO international pilot study of schizophrenia. *Archives of General Psychiatry*, **31**, 37–42.

TOR K. LARSEN, MD, Stavanger University Hospital, Division for Psychiatry, and Department of Psychiatry, University of Bergen, Norway; INGRID MELLE, MD, SVEIN FRIIS, MD, Ullevaal University Hospital, Norway; INGE JOA, RN, JAN OLAV JOHANNESSEN, MD, Stavanger University Hospital, Division for Psychiatry, Norway; STEIN OPJORDSMOEN, MD, ERIK SIMONSEN, MD, Roskilde County Psychiatric Hospital Fjorden, Roskilde, Denmark; PER VAGLUM, MD, Departments of Behavioural Sciences in Medicine, University of Oslo, Norway; THOMAS H. McGLASHAN, MD, Department of Psychiatry, Yale University, New Haven, Connecticut, USA

Correspondence: Tor K. Larsen, Stavanger University Hospital, Psychiatric Clinic, Armauer Hansensv. 20, pb 8100, N-4068 Stavanger, Norway. Email: tklarsen@online.no