Fatal and non-fatal repetition of self-harm

Systematic review

DAVID OWENS, JUDITH HORROCKS and ALLAN HOUSE

Background  Non-fatal self-harm frequently leads to non-fatal repetition and sometimes to suicide. We need to quantify these two outcomes of self-harm to help us to develop and test effective interventions.

Aims  To estimate rates of fatal and non-fatal repetition of self-harm.

Method  A systematic review of published follow-up data, from observational and experimental studies. Four electronic databases were searched and 90 studies met the inclusion criteria.

Results  Eighty per cent of studies found were undertaken in Europe, over one-third in the UK. Median proportions for repetition 1 year later were: 16% non-fatal and 2% fatal; after more than 9 years, around 7% of patients had died by suicide. The UK studies found particularly low rates of subsequent suicide.

Conclusions  After 1 year, non-fatal repetition rates are around 15%. The strong connection between self-harm and later suicide lies somewhere between 0.5% and 2% after 1 year and above 5% after 9 years. Suicide risk among self-harm patients is hundreds of times higher than in the general population.

Declarations of interest  JH's salary is paid by the mental health charity Leeds Mind from a research grant awarded by the UK National Lottery Charities Board.

We estimate that around a quarter of suicides are preceded by non-fatal self-harm in the previous year (Owens & House, 1994). If so, an episode of self-harm ranks with recent discharge from in-patient psychiatric care as the major risk factor for suicide (Gunnell & Frankel, 1994). This estimate of the link between self-harm and suicide needs to be accurate if we are to plan services aimed at reduction in suicide rate—a governmental priority for health improvement in the UK over recent years (Department of Health, 1999; Secretary of State for Health, 1999) and the target of a recent initiative by the USA Surgeon General (Vastag, 2001). Suicide is, nevertheless, too infrequent to be the main outcome event for a clinical trial of intervention after non-fatal self-harm. Instead, trials will continue to be designed to determine whether an intervention reduces the non-fatal repetition rate. Consequently, reliable estimates of repetition rate are needed for power calculation. We have undertaken a systematic review of the published literature in order to produce the best available estimates of rates of subsequent suicide and of non-fatal repetition following self-harm.

METHOD

Search strategies for the four databases Cinahl, Embase, Medline and PsycLit (each searched from their earliest entries) were constructed in 1998 for a non-systematic review (NHS Centre for Reviews and Dissemination, 1998) by an expert database searcher at the UK National Health Service Centre for Reviews and Dissemination, in conjunction with our clinical research team. We updated the strategies and ran them again in April 2001 for the present review. Ten journals were hand-searched for the Cochrane review of self-harm treatment trials (Hawton et al, 2001) but no extra hand-searching was carried out for the present review.

From the primary studies and all their secondary references, we included in our review every research report that fulfilled four criteria. The studies we selected were written in English, were published after 1970, described patients recruited to a study after attending a general hospital as a result of an episode of non-fatal self-harm and reported the proportion that repeated self-harm—fatally or not—for any follow-up period of at least a year. Suicides in most primary studies included those that were definite (by verdict of a coroner or equivalent authority) or probable (open verdicts or equivalent judgement); definitions were too variable for us to discriminate further and we have included them all and used the above broad definition of suicide. Because our search strategy found only one small study from the Far East that met the above criteria, we excluded it; the final list consequently represents research from Europe, North America and Australasia.

We excluded studies where the sample was restricted to participants who were young or elderly or had a learning disability. We did not exclude primary studies whose subjects were selected according to some measure of severity, such as established multiple repetition of self-harm or attending for the first time. Instead, we combined all the data and then applied a quality scale (described below). The majority of the studies were observational in design. Where we used data from clinical trials we combined data from both treatment groups, because the Cochrane review of trials of self-harm management (Hawton et al, 2001) found no clear difference between outcomes for experimental interventions compared with treatment as usual. Where more than one published paper set out findings for the same sample, we extracted results from the most complete version.

Measuring the quality of the primary study findings

For each study reporting a 1-year rate of non-fatal repetition or suicide we applied a ten-point quality scale based on features of the method and analysis (Table 1).

Study size

We weighted the quality score in favour of larger studies because they estimate outcome with the greatest precision. Clinical trials tend to score low in these ratings because of small sample size. We previously...
found (NHS Centre for Reviews and Dissemination, 1998) that, for the studies reporting repetition of non-fatal self-harm within 1 year, the median proportion repeating was 16%. A follow-up study of 200 subjects (n=200) would generate a 95% confidence interval of 11–21% (or 16 ± 5%) around a sample estimate of 16% (Gardner et al, 1989). A more precise estimate can be derived from n=600: 13–19% (or 16 ± 3%). Because suicide is a rare outcome event, large sample sizes are needed for precise estimates. In the same way, we used the median from our previous review (3% suicide at 1–4 years of follow-up) to determine reasonably precise and achievable estimates: n=500 would generate a 95% confidence interval of 1.5–4.5% (or 3 ± 1.5%); n=950 provides a more precise estimate of approximately 2–4% (or 3 ± 1%).

Because suicide is a rare outcome event, large sample sizes are needed for precise estimates. In the same way, we used the median from our previous review (3% suicide at 1–4 years of follow-up) to determine reasonably precise and achievable estimates: n=500 would generate a 95% confidence interval of 1.5–4.5% (or 3 ± 1.5%); n=950 provides a more precise estimate of approximately 2–4% (or 3 ± 1%).

Study sample
All hospitals discharge home a substantial proportion of patients attending as a consequence of self-harm (Owens, 1990), which is as many as two-thirds from some accident and emergency departments (Kapur et al, 1998). Comprehensive studies of hospital contact therefore identify subjects at accident and emergency or equivalent walk-in or emergency departments at general or psychiatric hospitals. The next best procedure is to ensure that all cases admitted as in-patients are included. Weak designs use convenience samples such as lists of weekday routine referrals to the self-harm assessment service; there will be exclusion biases but it is not clear what they might be. The most obvious biases of all occur when studies confine their sample to mild or to severe cases, perhaps to first-time or to multiple-repeat patients. We awarded up to four points for sampling (see Table 1); the final score is a cumulative one according to the absence of noticeable bias. Clinical trials usually had numerous exclusions and tended to score low.

Ascertainment of outcome
We found that the studies determined subsequent suicides by one or more of three methods: by inspection of local coroners’ (or equivalent) records, looking for the names of the study subjects; by efforts to determine the whereabouts of each patient, for example using hospitals, general practitioners and their records; and by checking names and other personal details against national registration of deaths. The first of these methods is weak – missing those who move home, even by only a short distance, and those who change their names. We awarded a point each for use of the two better methods.

Non-fatal repetition is more difficult to determine because of inadequate collection of data in most hospitals. We awarded half a point each for four steps taken to maximise identification of all the repeat episodes: use of a catchment area for the inclusion of subjects; interview follow-up of subjects; checks in general practice records; and checking of accident and emergency records.

Analysis of data
Many studies wrongly estimated the proportion repeating by recruiting subjects over a long period and following them up to a single end-point, failing to correct for the difference between subjects in the time-period denominator. Where a study used a uniform follow-up period – for example, everyone followed up for exactly 1 year from the date of inclusion – we awarded a point. Studies that used survival analysis scored a further point.

Combining the studies into a summary
The studies emerging from the literature search included single group cohorts, cohort analytical studies and clinical trials. This body of research is too heterogeneous for meta-analysis (Egger et al, 1998). Instead, we have placed the findings in rank order and we report their medians together with their interquartile range (25th–75th centiles).

RESULTS
The search strategy identified 90 studies meeting our inclusion criteria. Studies from the UK and Ireland accounted for over one-third (36%) of all the investigations. The others were undertaken in Scandinavia and Finland (26%), the rest of Europe (19%), North America (11%) and Australia and New Zealand (8%).

The main results of our analysis, grouped by duration of follow-up, are shown in Fig. 1. The median proportion repeating non-fatal self-harm is 16% at 1 year and 23% in studies lasting longer than 4 years. For subsequent suicide, the increment in the median after a longer

| Table 1 Non-fatal repetition and suicide: quality scores for study estimates |
|-----------------------------|-----------------------------|
| Size                        |                           |
| n=200 or more               | 1                          |
| n=600 or more               | 1                          |
| n=500 or more               | 1                          |
| n=950 or more               | 1                          |
| Sample                      |                           |
| No obvious bias to mild or severe cases | 1 |
| No deliberate exclusions    | 1                          |
| All admitted cases included | 1                          |
| Accident and emergency sample | 1                         |
| Ascertainment of outcome    |                           |
| Individual subjects followed up (90% or more) | 1 |
| National death records consulted | 1 |
| Catchment area targeted     | 1/2                        |
| Subjects interviewed (80% or more) | 1/2                      |
| General practitioner records consulted (80% or more) | 1/2 |
| Accident and emergency records checked | 1/2 |
| Analysis of data            |                           |
| Proper denominator (uniform time or at-risk period) | 1 |
| Survival methods with censorship | 1 |
| Total                       | 10                         | 10
follow-up is relatively much more – from less than 2% at 1 year up to nearly four times greater in the studies lasting over 9 years.

**Subgroup analyses**

For repetition at 1 year and suicide at 1 year we rank-ordered the studies according to date of publication and compared the findings of the more recent and older halves (Figs 2a and 3a). Medians were largely unaffected by the split but there was a wider dispersion of values among the studies in the past 10 years.

The high proportion of studies from the UK led us to examine the 1-year findings according to whether studies were UK-based or from elsewhere (Figs 2b and 3b). For repetition, UK studies showed the same median values as the rest of the literature but were more narrowly grouped around that median. For the 1-year suicide rate, both the UK and other studies showed tight bunching but UK studies had a median nearly five times lower than that of the rest of the literature (Mann–Whitney W = 54.5, P < 0.001).

The comparisons of 1-year findings based on the quality scores of the primary studies are shown in Figs 2c and 3c. For repetition and then for suicide we placed the studies in rank order according to quality score and then compared the better findings (those above the whole-group median score) with those below the median. For repetition, the values for the better-quality findings bunch tightly around 15% (a similar median to the one we found for all 37 studies); for the poorer-quality findings, the values are more dispersed around a higher median (21%). Examining suicide, we find a similar pattern: the higher-quality findings are tightly grouped around a median (1.8%) identical to that of the whole group of 26 studies, and the poorer-quality findings are far more widely dispersed around a slightly higher median.

Figure 4 shows a larger proportion of high-quality findings among the reports of non-fatal repetition than among the reports of subsequent suicide. We might have predicted this disparity because we were aware of few large studies that could estimate suicide with precision.

**DISCUSSION**

**Systematic reviewing of observational research**

Search strategies and safeguards against publication bias are less well developed for reviews of observational studies than they are for clinical trials. Although we are likely to have missed studies from our review, the tight clustering around the medians in higher-quality studies indicates that we would have to unearth many good studies with findings in one direction before medians for repetition or suicide would shift very far.

We were struck by the relative absence of studies from the USA, in line with the few American studies about intervention following self-harm (Hawton et al, 2001). Publication bias seems an unlikely explanation; our search terms used standard procedures, and three of the four bibliographical databases that we used are American and thereby likely to bias in favour of American studies. Clinical epidemiological study of self-harm is uncommon in the USA, despite the huge scale of self-harm there (Vastag, 2001).

**Summary of quantitative findings**

Summing up our findings, it seems that a reasonable estimate of non-fatal repetition is 15–16% at 1 year with a slow rise to
Quality scores in the suicide studies were generally low, with a median quality score for all 26 studies of only 2.5 out of 10 (interquartile range 2–5). Scores for the 9 UK studies were not noticeably different from those of the 17 non-UK studies: UK study median quality score was 2 (2–5.5) and non-UK median was 3 (1–5), a difference without statistical significance (Mann–Whitney $W = 212$, $P = 0.6$).

We checked whether health service differences between the UK and elsewhere might have led the UK studies to concentrate on accident and emergency departments, thereby biasing their samples towards less severe episodes that result in discharge from accident and emergency. In 2 out of 9 UK studies and 4 out of 17 studies from other countries, the researchers followed up all the patients who attended, not just the admitted patients. Similarly, we found the same median scores for sampling (out of a maximum of four) in UK and non-UK studies: zero for each group, with the same upper quartiles of 3.5. We therefore found no evidence of a group difference based on differential attention to patients attending hospital and leaving without in-patient admission.

Consequences of the inconsistent findings about suicide

Although our review might suggest that suicide following self-harm has a substantially lower incidence in the UK than elsewhere, the cumulative findings about suicide after self-harm are too flimsy to rely on. We need to understand the links between non-fatal self-harm and suicide if we are to plan clinical services and intervention research properly. The best current UK estimate of hospital attendance due to self-harm is around 400 per 100 000 (Hawton et al., 1997); 0.5% incidence of suicide in the
next year after self-harm (our median estimate for UK studies) accounts for 2 per 100 000 population, which is one-fifth of the England and Wales suicide rate of 10 per 100 000. If the same calculation is applied to our 1.8% median estimate from the better-quality studies, then around two-thirds of suicides (7 per 100 000) might be preceded by non-fatal self-harm in the preceding year.

Whichever estimate is the closer to the truth, it is plain that national suicide prevention strategies ought to be based on up-to-date research into non-fatal self-harm. High-quality follow-up studies of self-harm will help to keep those strategies relevant to clinical needs. The studies that ought to be undertaken will be large, following up well over 1000 self-harm patients, and they will be based on all patients attending hospital, regardless of whether or not they were admitted from accident and emergency. Determining the outcome of those who are treated only in primary care will be feasible only when there is an increase in data-sharing in primary care. Repetition will be ascertained from accident and emergency or other hospital contact records, rather than from ward, special unit or discharge data. Suicides will be determined by the use of national records of the registration of deaths. The study data will be analysed using the statistical techniques of survival analysis.

Suicide is a rare event occurring in 1 in 10 000 people a year, and bringing about a reduction in the population’s suicide rates is a difficult challenge. Recent non-fatal self-harm indicates a large increase in individual risk – it is probably the major risk factor – but the incidence among these people rises to around 1%. Unfortunately, all our clinical methods for predicting suicide among our patients have a very poor positive predictive value at this low level of incidence (Geddes, 1999). Only a population strategy (Rose, 1992) is likely to achieve a reduction in the suicidal potential after self-harm – through application of an intervention aimed at all self-harm patients. But current evidence tells us that the few clinical trials of intervention after self-harm are characterised by inadequate power, unrepresentative samples and unsuitable data analysis (Hawton et al., 1998). The second research need is therefore for the first-ever large, well-designed clinical trial of brief intervention after non-fatal self-harm.

ACKNOWLEDGEMENTS

We thank Julie Glanville of the NHS Centre for Reviews and Dissemination at the University of York for the first systematic search and Lesley Patchett of the School of Medicine at the University of Leeds for locating and organising the studies.

APPENDIX — STUDIES REVIEWED


Deliberate self-harm: a follow-up study of 279 patients.


McFarland, B. H. & Beavers, D. J. (1986)


Prospective study of fluoxetine treatment and suicidal therapy vs insight-oriented therapy for repeated suicide attempts.

Spinal Cord


1060-1064.

Efficacy of different strategies of

American Osteopathic Association


Efficacy of different strategies of aftercare for patients who have attempted suicide.

Journal of the Royal Society of Medicine, 90. 496–498.

McFarland, B. H. & Beavers, D. J. (1986)


McEvedy, C. J. (1997)


Journal of the Royal Society of Medicine, 90. 496–498.

Nordentoft, M., Brem, L., Munck, L. K., et al. (1993)

High mortality by natural and unnatural causes: a 10 year follow up study of patients admitted to a poisoning treatment centre after suicide attempts. BMJ. 306, 1637–1641.


Frequent deliberate self-harm: repetition, suicide and cost after three or more years. Scottish Medical Journal, 40, 10–12.


The serious suicide attempt: epidemiological and follow-up study of 886 patients.


Rosenman, S. J. (1983)


CLINICAL IMPLICATIONS

- The link between self-harm and suicide is a strong one; subsequent suicide occurs in somewhere between 1 in 200 and 1 in 40 self-harm patients in the first year of follow-up and in around 1 in 15 people after 9 or more years.

- Non-fatal repetition is common after self-harm; about one in six patients repeats over the next year and one in four after 4 years.

- The UK estimates of rates of suicide after self-harm are low when they are compared with the rest of the research literature.

LIMITATIONS

- Estimates of fatal and non-fatal repetition after self-harm are derived from an accumulation of small studies rather than from large-scale monitoring.

- Estimates of rates of subsequent suicide are largely derived from poor follow-up data.

- Pooled estimates of subsequent suicide are therefore imprecise.

---

Owens ET AL

---


REFERENCES


FATAL AND NON-FATAL REPETITION OF SELF-HARM


