How Many Account for How Much? Concentration of High-Risk Sexual Behaviour Among Gay Men

Anthony P. M. Coxon
University of Essex, England

Thomas J. McManus
The Royal London Hospital, England

The data set of 2,182 gay/bisexual men's month-long sexual diaries collected by the United Kingdom Socio-sexual Investigations of Gay Men and Aids (Project SIGMA) was used to analyse the extent to which acts of anal intercourse (AI) are distributed among gay men, using the Lorenz/Gini methodology for measuring concentration. Most individuals (60%) who engage in AI do so only once or twice a month, but there is a long tail of those who do it much more. In terms of the amount of AI acts, one tenth of the individuals are performing half of the acts of AI. The Gini coefficient of concentration is high (0.55). Factors most affecting rates and concentration of risk behaviour include relationship status, HIV sero-status and concordant/discordant partner status.

The number of individuals who engage in risk-prone sexual behaviour is of obvious and primary importance in the analysis of AIDS/HIV infection. Anderson (1988) points out that the degree to which this infection will spread is basically dependent on the magnitude of the basic reproductive rate, but that a small proportion of individuals with high rates of partner change (and hence of risk-prone sexual behaviour) make a disproportionate contribution to the reproductive potential of the infection. Hence, the amount of such risk activity is equally as important as the number engaging in such activity and, taken together, they define the concentration of risk-prone sexual behaviour.

One reason why concentration of sexual behaviour has received relatively little attention in the literature is due to the difficulty of obtaining reliable and accurate measures of the actual amount of risk-prone sexual behaviour (at least by the use of conventional methods of data collection). As an antidote to recall bias in surveys (Baddeley, 1979; Coxon 1993a, 1993b; Dex, 1991), self-completion methods have been proposed which record the events simultaneously, or only a short period after they happened. In particular, diary methods have been used successfully for studying sexual behaviour in general (Fortenberry, Cecil, Zinet, & Orr 1997; Freeman, DeRubeis, & Rickels 1996; Graham & Bancroft 1997; James, Bignell, & Gillies 1991; Leigh, Gillmore, & Morrison 1998; Reading, 1983) and in studies of gay men in particular (Coxon, 1994, 1996; Coxon & Coxon, 1993a, 1993b; Gold & Rosenthal, 1995; McLawns et al., 1990). Moreover, diaries have been found to produce data which are as reliable as, and in most cases more valid than, questionnaire procedures (Conrath, Higgins, & McLean 1993; Coxon, 1996, 1999; Freeman et al., 1996; James et al., 1991; Juster, 1985; Leigh et al., 1998; McLawns et al., 1990; Phellas, 1994; Poikolainen & Karkkainen, 1993; Robinson, 1985; Weinhardt, Forsyth, Carey, Jaworski, & Durant, 1998).

In this paper we rely upon the data derived from month-long sexual diaries collected by Project SIGMA in England and Wales (Coxon 1996) in order to study various types of homosexual anal intercourse (AI) and to do three things:

1. To integrate information on incidence (numbers of men engaging in AI) and rates (amount of AI being done) by means of the Lorenz Diagram, and to examine the concentration of such risk behaviour using the associated Gini coefficient.

2. To see if different variants of high-risk behaviour (e.g., protected vs. unprotected anal intercourse) differ in concentration.

3. To see what variables affect concentration.

METHOD

Research Instrument: The Sexual Diary

The method of sexual diaries (MSD), was developed and used within Project SIGMA from 1984.¹ Diaries are written in natural English (but according to a structured schema),

1The sexual diary method has been developed in conjunction with interview and serological testing of gay and bisexual men in England and Wales, by Project SIGMA (Coxon, 1988, 1994, 1996, 1999).
and are filled in on a daily basis. The method yields highly contextualised and detailed information of the sequence of sexual activities (Parker & Carballo 1990). The diary format is based on the unit of the sexual session (consisting of a sequence of sexual acts with a natural time-marked start and end), in turn consisting of sexual acts, each of which includes specific information on agency (what we call modality—who does what to whom), on the sexual behaviour itself (e.g., anal intercourse), and on the precise destination of the ejaculation of semen (if it happens) with respect to both partners.

**Data Set**

The sexual diary data-set used in this paper consisted of the 2,182 individual diary/months of 1,035 men who have sex with men, drawn from ten locations/sites in England and Wales in the four waves of Project SIGMA between 1987 and 1992. Most of the analysis focussed on the subset of 628 diary/months which included one or more acts of anal intercourse in this period.

**Measures**

The three variables of interest are:

1. The amount of risk-activity, measured by the number of sessions which contain one or more risk-acts (identified and extracted from the diary record);

2. The number of men engaging in a given risk-activity such as anal intercourse. Information is obtained by counting the number of diarists who engaged in the activity during the standard month period, and is also extracted from the diary-record itself; and

3. The riskiness of a sexual session, measured in terms of what combination of risk-factors occurred. The risk factors measured were whether anal intercourse (AI) occurs in a session, whether ejaculation (E) occurs as a result of the act of AI, and whether a condom is worn in the act of anal intercourse.

A typology of increasing risk-activity is presented in Table 1. This simple typology differentiates risk sessions (i.e., those involving AI) in terms of whether ejaculation occurs and, if so, whether or not it is protected. In brief, high-risk sessions are any which involve ejaculation, and highest-risk sessions are differentiated in terms of whether or not the ejaculate is into a condom (C: protected) or into the receptive partner himself (D: unprotected).

### Table 1. Typology of Risk Sessions

<table>
<thead>
<tr>
<th>Type</th>
<th>Anal</th>
<th>Intercourse</th>
<th>Ejaculation</th>
<th>Condom</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>B</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>C</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>D</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>E</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
</tr>
</tbody>
</table>

Legend: ✓ must contain
* may or may not contain
× must not contain

Concentration of ownership in the distribution of resources in general, and of income, wealth, and property in particular. Subsequently, it has been extended to deal with inequality and concentration in other domains, such as political power and resources (Alker, 1965), geographic and regional concentration (Shelburne & Bednazl, 1993; Smith, 1979), and health (Davey Smith & Egger, 1996). In this paper, the concentration of sexual risk activity is measured using the same methodology.

The Lorenz Diagram gives a visual representation of inequality and concentration in a single graph. Its use in the context of sexual behaviour can be illustrated by reference to the following (fictitious) set of data (see Figure 1). Suppose we have found the following empirical information about the distribution of sexual acts (having sorted the

![Figure 1. Example of a Lorenz Curve (fictitious data).](image-url)

---

1. A core of 180 men provided a month-diary in each of the four waves; 413 provided one diary, and the remaining 442 men average 2.24 diaries each.
2. The hand-written natural-language diaries (in anonymised micro-fiche form) are archived in the Contemporary Medical Archives Centre of the Wellcome Institute, London. The diaries are also encoded into a database format, where the sexual session is the record. Special-purpose software, SDA: Sexual Diary Analysis (Coxon & Coxon, 1993b) then extracts information and counts from the records. The machine-readable data in data-base form and the software for their analysis are available at cost from the author.
3. The analysis is done using the LORENZ program in the SDA suite of programs (Coxon & Coxon, 1993b).
data by those who do exactly 1 sex act, exactly 2 acts, and so on): The first 50% of the men do 18% of the sex acts, the first 75% of the men do 30% of the sex acts, the first 90% of the men do 50% of the sex acts, and 100% of the men do 100% of the sex acts. This provides the basic information which the Lorenz diagram portrays (see Figure 2). The diagram is designed to display what proportion of the population "owns" what proportion of the "scarce resource" by charting what proportion of the population (horizontal axis) has engaged in what proportion of the total number of sexual acts (vertical axis). The data are plotted for 1 act, up to 2 acts, up to 3 acts, and so on, and the Lorenz Curve (guaranteed by its cumulative nature to be monotonic upward), formed by joining the points, is portrayed in Figure 1.

This Lorenz curve is then compared to what would be the case if sexual acts were shared equally, so that 10% of the population did 10% of the sexual acts, 20% did 20% of the sexual acts, and so on. This forms the line of equality (or equal distribution) moving up diagonally at 45° from the bottom left to the top right of the chart, portraying the baseline "equality" distribution on what would be the case if the amount of sexual activity were the same for all. The difference between these two curves gives a visual representation of the relative inequality, or of concentration—how far from equality the distribution it is, and how far the sexual activity is concentrated in a small proportion of individuals. A numerical measure of overall inequality or concentration is given by the Gini coefficient, which measures the gap between equal distribution and actual distribution. The value of the Gini coefficient is 1 where there is total concentration/inequality (the very topmost fraction owns everything) and 0 if there is total equality or equal dispersion (everybody owns the same amount).

The Lorenz curve and associated Gini coefficient thus can be used in this case to measure how the stock of (risky) sexual activity is divided between the gay men involved, and can provide a useful numerical measure of the concentration of sexual risk. The substantive questions that can be addressed by these means are as follows:

1. What is the distribution over the specified levels of risk activity?

2. Are individuals equally liable to engage in unprotected variants? Or is high-risk sex concentrated in relatively few individuals? If so,

3. If high-risk sex is concentrated in relatively few individuals, what characteristics/contexts typify them?

---

5 Strictly, twice this area for the whole diagram. For a continuous variable, the Gini coefficient is related to the Lorenz curve L by:

\[ G = 1 - 2 \int_0^1 L(u) \, du \]

and for a discrete variable with p categories sorted in order from the lowest to the highest, the coefficient is defined as:

\[ G = \sum_{i=1}^{p} \frac{X_i - Y_i}{2} \Delta X_i \]

where \( X_i \) is the cumulative percentage of X by unit, and \( \Delta X_i = X_i - X_{i-1} \)

See also Alker (1965, pp. 41-42).

---

Figure 2. Outcomes of acts of anal intercourse (from sexual diary data).

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Out of the 2107 (=100%) acts of anal intercourse (Potential Risk)</td>
<td></td>
</tr>
<tr>
<td>30% did not result in ejaculation of semen ([N]), and</td>
<td></td>
</tr>
<tr>
<td>70% involve ejaculation of semen. (Potential High Risk)</td>
<td></td>
</tr>
<tr>
<td>Of these ejaculations,</td>
<td></td>
</tr>
<tr>
<td>17% of ejaculations are into a condom ([C]) (HR Protected)</td>
<td></td>
</tr>
<tr>
<td>but 53% are unprotected. (HR Unprotected)</td>
<td></td>
</tr>
<tr>
<td>Of these unprotected ejaculations,</td>
<td></td>
</tr>
<tr>
<td>12% go &quot;elsewhere&quot; ([E])</td>
<td></td>
</tr>
<tr>
<td>4% go on the partner's body ([O/I])</td>
<td></td>
</tr>
<tr>
<td>37% go into the partner's anus ([H/M]) (Highest Risk)</td>
<td></td>
</tr>
</tbody>
</table>

---

RESULTS

Analysis is restricted to data from those men who engaged in anal intercourse during the month of their sexual diary (one third of the total).6

Amount of Anal Intercourse

First we examined the information about the amount, volume, or "stock" of the 2,481 acts of AI, without reference to the number of individuals involved. The outcomes are reported in Figure 2 as a Risk Hierarchy. This shows that if an act of anal intercourse occurs, it is most likely (70% of acts) to result in ejaculation; if ejaculation occurs, it is more likely to be without a condom (53% of all AI acts) than with a condom; if ejaculation occurs, it is most likely to be directly into the partner's anus (37% of all AI acts), than elsewhere; and if ejaculation is into the partner's anus, it is more than twice as likely to be without a condom (37% of all AI acts) than with a condom (17% of all AI acts). However viewed, these figures indicate that in this sample, anal intercourse to ejaculation without a condom into the partner's anus (AI/NC) is the rule rather than the exception in gay men's sexual intercourse.

Concentration of Risk

AI Baseline. Given this extent of highest risk behaviour, the question arises of how risk behaviour is concentrated, and whether the concentration increases as acts become more high-risk. To do this, three levels of the typology of Table 1 are considered: B: High Risk Session—anal

---

6 In the 5-wave diary data set, two thirds of these (68%) contained no act of anal intercourse during the month concerned. The subset of interest is now the 628 (diary/months) which included one or more acts of anal intercourse.
intercourse (AI); C: Highest Risk Session (Protected)—ejaculation into a condom (AIE/C); and D: Highest Risk Session Unprotected—ejaculation into anus (AIE/NC). In so doing, attention can be directed to two main questions: Is high-risk sex concentrated in a few individuals or is it evenly spread, and are there characteristics which distinguish, say, the occasional practitioner from those for whom anal intercourse is the central part of their sexual repertoire? The basic information is given in Table 2.

Thus, 39% do only one act of AI per month, and this accounts for 10% of the stock of all 2,481 acts of anal intercourse. The distribution of men who do exactly $p$ acts of anal intercourse in a given month (Table 2, col. 2) is highly skewed toward the low end, with a median of 2 acts/month (mean of 4), and a long tail, which peters out to one man who engages in 104 acts in a month. The distribution of the number of acts (Table 2, col. 6) is quite different, and shows clearly that the bulk of acts of anal intercourse is concentrated in the higher reaches: A few men are accounting for a high proportion of the acts. These two constituent percentage distributions are presented (in the same scale) in Figure 3.

Figure 3a. Percentage of individuals who do exactly $x$ acts of AI a month. Results from SIGMA diary data.

Figure 3b. Percentage of all acts of AI accounted for by those who do exactly $x$ acts a month. Results from SIGMA diary data.
In Figure 3a, the first bar gives the percentage of the individuals who did exactly one act of AI a month, the second gives the percentage of those who did exactly two acts a month, and so on. The distribution is massively skewed and is dominated by those who engaged in AI only a few times a month, with a long tail consisting of the very few who engaged in a large amount of AI. Figure 3b presents the percentage of the acts that were accounted for by those who only engaged in AI once a month, twice a month, and so on, up to the amount accounted for by the individuals who engaged in AI a very large number of times. This distribution is also skewed, but is far less dominated by those who engaged in AI relatively rarely, and indicates how the few "extremists" account for considerable fractions of the acts of AI.

These two distributions are now put together in the Lorenz diagram in Figure 4. As the number of AIs/month goes up, the cumulative percentage of men (Table 2, col. 4) increases more rapidly than the increase in the cumulative percentage of AIs which represents (Table 2, col. 7). So, when 10 AIs/month is reached, a full 93% of men but only 60% of the AIs have been accounted for. From 10 AIs up, the percentage of men unaccounted for dwindles, and beyond 20 AIs/month only 16 (2.3%) are left—but between them, this 2.3% of men are responsible for 23% of the AIs. Hence, the few in the higher reaches of the tail of the distribution contributed very disproportionately to the amount of anal intercourse: The top 10% of men account for almost half of the acts.7

The Gini coefficient for these data is 0.55, which is high on any account, comparing values of 0.20 to 0.45 for income inequality in Western democracies and 0.45 and 0.60 for developing countries (Deininger & Squire, 1996). A coefficient value of this size indicates that there is a high degree of concentration of potentially high-risk behaviour in a small fraction of gay men (or, equivalently, that anal intercourse is distributed in a highly unequal manner); moreover, anal intercourse is more highly concentrated than other sexual activity among gay men.8

This suggests that, with respect to concentration of anal intercourse, there are three groups of gay/bisexual men: (a) those who do not engage in anal intercourse, (b) those who do engage in a small amount of anal intercourse, and (c) a small group accounting for a high fraction of the activity.

How true is this? We know already that there is a sizeable group (about one third) who does not engage in AI, and there is also a sizeable group who does engage in AI, but usually between one and three times a month—less than once a week. For the first two groups, then, the description is basically accurate. The third group is more diffuse: There are certainly some men who engage in AI very frequently and who account for a significant fraction of the potential risk acts, but any exact point of division is bound to be arbitrary. If we take the top tenth, they may account for half of the acts of anal intercourse, but they range from the 2% who engage in AI 8 times a month, to the 0.15% who engage in AI 40 times a month and the one man who engages in AI 104 times a month. Those at the extreme end of the distribution cannot be dismissed as an amorphous, dispersed remnant. The sexual activity of even a few such individuals can have major epidemiological consequences for the transmission of HIV if they are HIV-antibody positive, if their anal intercourse is unprotected, or if they are primarily insertive partners. Are they "risk-rich"? One important clue9 is that those in the highest quartile are considerably more likely to carry a condom with them at all times (perhaps because they expect to engage in AI) compared to the lower quartile, who do not, and for whom the relatively rare AI "catches them by surprise," unprepared and unprotected.

Variants of Risk Acts

Does concentration increase for the variants of AI which are more directly implicated in HIV transmission? To answer this it was necessary to differentiate unprotected from protected anal intercourse, and to allow for differences arising from the modality (insertive/receptive role) of the act. Taken together, this defines four types of AIE, contrasting protected versus unprotected and insertive (active) versus receptive (passive) variants.10

In order to examine significant differences between these four types of AIE, it is appropriate to look at both the

---

7The full data show that the top [10%, 5%, 2%, and 1%] of men account for [46%, 33%, 20%, 15%] respectively of the acts of AI.
8G = 0.68 for all sexual acts (Coxon, 1996, pp. 191-192).
9This information is obtained from the SIGMA interview data.
10In the sexual diary these are defined as [AF, H*], [FF, *M], [AF, C*], and [PF, *C].
Table 3. Concentration Characteristics of Four Types of AIE

<table>
<thead>
<tr>
<th>Summary characteristics</th>
<th>Unprotected</th>
<th>Type of AIE</th>
<th>With condom</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Active</td>
<td>Passive</td>
<td>Active</td>
</tr>
<tr>
<td>No. of fuck-acts/month</td>
<td>434</td>
<td>442</td>
<td>362</td>
</tr>
<tr>
<td>No. of individuals</td>
<td>188</td>
<td>154</td>
<td>163</td>
</tr>
<tr>
<td>Average no. fuck-acts/month</td>
<td>2.31</td>
<td>2.87</td>
<td>2.22</td>
</tr>
<tr>
<td>Dispersion (IQR)</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>% who do exactly 1 fuck-act/month</td>
<td>54%</td>
<td>51%</td>
<td>55%</td>
</tr>
<tr>
<td>% who do 1 or 2 fuck-acts/month</td>
<td>75%</td>
<td>73%</td>
<td>75%</td>
</tr>
<tr>
<td>Top 1/10th of individuals account for n% of fuck-acts</td>
<td>34%</td>
<td>46%</td>
<td>32%</td>
</tr>
<tr>
<td>Gini coefficient</td>
<td>0.42</td>
<td>0.51</td>
<td>0.42</td>
</tr>
</tbody>
</table>

numbers involved (numbers of both acts of and of individuals) and also at the shape of the distributions—the level (average), spread (dispersion), and straggle (concentration at high end and low end). The summary information is given in Table 3. The most astonishing thing about Table 3 is how similar the four types of highest risk behaviour are to each other, and also how much the characteristics differ from those that hold for all the acts of anal intercourse taken together. In particular, we found that:

1. The average number of unprotected acts of anal intercourse is considerably smaller than the overall average for all acts of anal intercourse.
2. The overall spread of unprotected distributions is smaller, showing a more compact form and concentration.
3. The low-end concentration (the extent to which the highest-risk anal intercourse behaviour is concentrated in the 1-a-month, or at most the 2-a-month, category) is very marked and is a good deal more concentrated at this low level than general anal intercourse is.
4. The high-end concentration (the share of the acts of anal intercourse of the top 10% of individuals) is very considerably lower for the highest-risk acts (about 35% versus 54% for all anal intercourse). So although there are some individuals who have a disproportionate share in acts of unprotected anal intercourse, this share is a lot less than for the overall rate of anal intercourse.

Perhaps more unexpected is the fact that using a condom makes little difference to the characteristics and concentration of the distributions. The differences in high-risk behaviour due to modality is even less marked, except that receptive variants of anal intercourse seem slightly more common. However, this is affected by one extreme outlier who engaged in receptive anal intercourse on an average of twice a day in the month (thereby accounting by his own activity for 14% of the highest-risk sex). The major characteristic of highest-risk acts of anal intercourse, whether protected by a condom or not, is that most of them (well over half) are concentrated in those who only engage in one act per month. With that information, it becomes clear that in terms of number of men involved, highest-risk acts of anal intercourse are concentrated primarily among those who only engage infrequently, and this merits attention.

Variables Affecting Risk And Its Concentration

Finally, it is relevant to summarize the results of searching for pockets of "risk-rich" gay men among the sexual diarists. The three most potent differentiating variables are: relationship status (closed/open/no regular relationship); HIV sero-status (tested positive/last test negative/not tested) and concordant-discordant partner status.

Relationship status. Closed relationships had a high amount of unprotected AIE, and 14% of those in an ostensibly closed relationship engaged in unprotected anal intercourse at least once in a month with another partner. In open relationships there was less unprotected AIE with a primary partner and more high-risk sex with other partners. Those with no regular relationship tended to avoid acts of AIE, make more use of condoms, and ejaculate more "on" than "in" a partner. Hence, highest risk AIE is concentrated with the primary partner for the closed relationships, and with other partners for those in an open relationship.

HIV sero-status. HIV-positive men have a quite different profile than men testing negative. Positive men had more sex, had more anal sex, and had more anal sex to ejaculation than the negative men — but positive men had considerably more protected AIE, and engaged in a higher degree of receptive AIE than negative men. Those not tested make less use of condoms than either other group.

Concordant/discordant partner status. Concordant highest-risk sexual activity (AIE/NC) (whether concordant positive or concordant negative) is far more prevalent than dis-
cordant AIE (30% of sessions are negative concordant and 26% are positive concordant vs. only 2% which are discordant sessions). But sessions where the diast is negative and where his partner is either untested or his HIV status is unknown tend to involve the highest rates of highest-risk sex. Here, probably, is the greatest reservoir of concern.

**DISCUSSION**

Behavioural interventions among gay men are predicated upon data findings that anal intercourse among gay men is a minority behaviour, that condom use differs depending on whether the partner is regular or casual, and that condom use is normally consistent with principles of risk-reduction. Most of these findings are broadly supported both by the SIGMA interviews (Davies et al., 1993) and diary studies (Coxon, 1996). If AI is viewed as a necessary condition for risk, then it is clear that it is highly concentrated: There is a significant number of individuals who are responsible for a high fraction of high-risk sex, and their sexual activity can have far-reaching consequences. But the concentration of highest-risk anal intercourse is primarily in the relatively infrequent acts of a relatively large number of gay men (rather than in the very frequent acts of a few)—and it is this which is likely to lead to more rapid diffusion of infection and, ultimately, to higher levels of infected individuals.

**REFERENCES**


Manuscript accepted June 9, 1999.