Effect of request for unique personal identifiers and souvenir incentives on consent to health record linkage: evidence from an RCT nested within a cohort

Michael Y. Ni¹, Tom K. Li¹, Rex W.H. Hui², Ian McDowell³, Gabriel M. Leung¹

Affiliations:

1. School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong Special Administrative Region, China.

2. Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong Special Administrative Region, China.

3. School of Epidemiology, Public Health and Preventive Medicine, University of Ottawa, Ottawa, Canada

Corresponding author:

Dr. Michael Y. Ni, School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, People's Republic of China.

Tel: +852 3917 6776; Fax: +852 3520 1945; email: nimaryhkehk.

Word count (abstract): 218

Word count (main text): 1,598

Number of tables/figures: 6

Conflict of interest: All authors report no competing interests.

Running title: Effect of personal identifiers and incentives on health record linkage
ABSTRACT

Objective: It is unclear if unique personal identifiers should be requested from participants for health record linkage: this permits high-quality data linkage but at the potential cost of lower consent rates due to privacy concerns.

Study Design and Setting: Drawing from a sampling frame based on the FAMILY Cohort, using a 2x2 factorial design, we randomly assigned 1,200 participants to: (1) request for Hong Kong Identity Card number (HKID) or no request, and (2) receiving a souvenir incentive (valued at USD4) or no incentive. The primary outcome was consent to health record linkage. We also investigated associations between demographics, health status, and postal reminders with consent.

Results: Overall, we received signed consent forms from 33.3% (95% CI 30.6% to 36.0%) of respondents. We did not find an overall effect of requesting HKID (-4.3%, 95% CI -9.8% to 1.2%) or offering souvenir incentives (2.4%, 95% CI -3.1% to 7.9%) on consent to linkage. In subgroup analyses, requesting HKID significantly reduced consent among adults aged 18-44 years (OR 0.53, 95% CI 0.30 to 0.94, compared to no request). Souvenir incentives increased consent among women (OR 1.55, 95% CI 1.13 to 2.11, compared to no souvenirs).

Conclusions: Requesting a unique personal identifier or providing a souvenir incentive did not affect overall consent to health record linkage.

Keywords: Health record linkage; Data linkage; Consent; Randomised; Unique identifier; Incentive
What is new?

- Overall, requesting a unique personal identifier or providing a souvenir incentive did not significantly affect consent to health record linkage. Thus in this context, unique personal identifiers should be requested to enable higher precision in health record linkage.

- However, the effect of the interventions varied by age and sex: younger respondents were less likely to provide consent when asked to provide a unique personal identifier, whereas incentives increased consent among young people and women.

- Postal invitations may still be a relatively inexpensive, useful initial strategy to obtain consent for health record linkage.
INTRODUCTION

As participation rates continue to decline for academic research, government or industry surveys (1), data linkage to administrative and medical records provides an important alternative to access routinely collected data while minimising respondent burden (2). Yet low consent rates to health record linkage and response bias are frequent challenges (2-4). Age (5-9), ethnic minority status (6, 7, 10, 11), education level (5, 7, 8), household income (5, 6), area of residence (6, 10), health status (5, 7, 12), chronic metabolic diseases (5), depressive symptoms (12), and health services utilisation (7) have been associated with consent, although inconsistencies remain (7-9, 11, 12). In addition, while cash incentives appear to increase response to mail surveys or other survey modalities (13-16), it is unclear whether incentives influence consent for data linkage (17), and whether souvenir incentives are effective.

Data linkage can be achieved in two ways. Deterministic linkage uses a unique personal identifier - such as the Hong Kong Identity Card number (HKID), or the Social Security number in the US or the National Health Service number in the UK – to link information contained in different datasets. The advantage of a unique identifier is higher health record linkage success. However, HKID numbers are also used in banking services and as authentication for password changes, thus are generally regarded as sensitive personal data (18). Privacy concerns have been shown to be negatively associated with consent in observational studies (19, 20). The second option, probabilistic linkage, uses personal data such as name and date of birth to circumvent the need for unique personal identifiers but may result in duplicate matches. The choice is therefore unclear: unique personal identifiers permit high quality data linkage but at the potential cost of lower consent rates. We accordingly tested the effect of 1) requesting a unique
personal identifier and 2) providing a small, souvenir incentive on consent to health
record linkage, using a randomised factorial design. A secondary objective was to
identify whether respondent characteristics and the use of postal reminders were
related to consent.

METHODS

Study design and participants

We nested a randomised controlled trial within the FAMILY Cohort (total N = 46,001), a
population-based longitudinal study described in detail elsewhere (21). Participants
were drawn from a subsample of the FAMILY Cohort where one member from each
household was randomly selected to form the sampling frame. Eligibility was defined
by age ≥ 18 years and completion of two waves of in-person follow-ups. We used a 2-by-
2 factorial design to randomly assign 1,200 adult participants to: (1) request for
personal unique identifier (HKID) or no request, and (2) receive a souvenir incentive or
no incentive. Randomisation sequence was created using Stata MP 13.1 (StataCorp,
College Station, TX), and was stratified by sex, age group and educational attainment
(Figure 1). All participants were blinded to the randomised design. The study was
approved by the Institutional Review Board of the University of Hong Kong/Hospital
Authority Hong Kong West Cluster.

All respondents received identical invitation letters, an information leaflet, a consent
form and a prepaid return envelope by mail. For respondents allocated to receive HKID
requests, their consent form included a box to fill in their HKID and, as an alternative, a
secure web link to enter the HKID (Figure 1). Either was accepted. Those allocated to
receive a souvenir incentive were mailed an upfront souvenir including a FAMILY
Cohort-branded washcloth, a set of chopsticks and a rice paddle, with a collective value of USD~4. Two reminder letters were sent 20 days and 33 days after the initial mailing. We allowed 60 days as the cut-off period for return of completed consent forms.

**Primary analysis**

The primary outcome was receipt of written consent to health record linkage by intention-to-treat. We needed an overall sample size of 1,188 (alpha = 0.025 and power = 0.90) for the 2 factorial comparisons (hence, overall alpha = 0.05 and power = 0.81) in order to detect a 10% absolute difference in consent proportion between (1) HKID request vs. no request, and (2) souvenir incentive vs. no incentive (22). The expected consent proportion for the control group was 35%. We used the chi-square test to compare the proportions consenting between the four groups. We then used multivariable logistic regression and the likelihood ratio test to perform interaction analyses to test whether the effect of requesting HKID varied according to whether participants were randomly assigned to receive a souvenir incentive. Similarly, we performed tests for interactions between the interventions and demographic subgroups (age, sex, education level and household income) by adding treatment subgroup interaction terms to the models.

**Secondary analyses**

We assessed associations between baseline predictors assessed at wave 2 including demographics and health status (physical and mental well-being using the SF-12v2 (23, 24), chronic disease status and hospital admission) with consent using multivariable logistic regression with Bonferroni correction for multiple comparisons. Each baseline
predictor was examined in a separate regression model, adjusting for age, sex, education level, employment status and household income.

Additionally we examined whether the two postal reminders were associated with a significant change in the daily consent rates using segmented logistic regression (25). All analyses were done using R version 3.3.0.

RESULTS

Participants (n=1,200) were randomly allocated into four groups: 1) HKID request and incentive (n=292); 2) HKID request without incentive (n=306); 3) incentive without HKID request (n=303); 4) no incentive and no HKID request (n=299). Baseline characteristics were balanced in the four groups (Table 1).

Consent by group

The proportions consenting were, in descending order, 36.6% (95% CI 31.2%-42.3%) in the group with an incentive and without a HKID request, 34.1% (28.8%-39.8%) in the control group receiving no incentive and no HKID request; 32.2% (26.9%-37.9%) in the group receiving both incentive and request, and 30.1% (25.0%-35.5%) among the group receiving HKID request without an incentive. There was no interaction between the two interventions (P=0.96).

Effect of HKID request

Participants in the two groups that received an HKID request had a 4.3% lower absolute consent (31.1%, 95% CI 27.4%-35.0%) compared to those not asked to give their ID (35.4%, 95% CI 31.6%-39.3%); the difference was not significant (P=0.14).
Effect of souvenir incentive

Those who received a souvenir incentive had a 2.4% higher absolute consent proportion (34.5%, 95% CI 30.6%-38.4%) compared to those without the incentive (32.1%, 95% CI 28.4%-35.9%); this difference was also not significant (P=0.38).

Interaction between demographics and interventions

Subgroup analysis shows that the effect of requesting the HKID varied according to age, where requesting HKID reduced consent proportions among younger adults but not among middle-aged or older adults (Figure 2). Provision of a souvenir incentive increased consent proportions among young adults and women (Figure 3).

Demographic correlates of giving consent

In the overall sample, older age and higher household income were associated with consent to health records linkage (Table 2). However, only age remained significant after correction for multiple comparisons. Other demographic characteristics and indicators of health status were not associated with consent.

Effect of postal reminders

There were clear temporal associations of receiving completed consent forms with each of the two reminder mailings (Figure 4, upper panel). Seven breakpoints in the daily consent rate were identified over time (Figure 4, lower panel). The daily consent rate increased substantially after the third (Day 20) and the fifth (Day 32) breakpoints, which correspond to the first and the second postal reminders.
DISCUSSION

We did not find an overall effect of requesting a unique personal identifier on consent to health record linkage. However, younger respondents were less likely to provide consent when asked to provide HKID (Figure 2), which is consistent with younger individuals having more privacy concerns about health record linkage (8, 26).

Comparison of responders and non-responders showed that older age was associated with consent (Table 2). There were no systematic differences regarding other demographic characteristics or health status.

We also did not find an overall effect of providing a souvenir incentive on consent, although incentives increased consent among younger people and women (Figure 3). The effect of the interventions did not vary by socioeconomic status. Our overall null effect for incentives could be attributed to the use of souvenirs rather than monetary incentives (13-15, 27). In addition, the incentives may not have increased consent proportions as participants were drawn from a cohort that has previously received similar incentives over a number of years.

Possible reasons for the low consent rate in our study include the use of postal invitations. Postal invitations are less costly but yield lower response rates compared to face-to-face interviews (28). Our level of consent is comparable to previous studies using postal invitations for health record linkage (3, 4). Moreover, a multi-ethnic national cohort study found that individuals of Asian ethnicity are less likely to consent with health records linkage (10). However, the study also identified higher educational attainment to be negatively associated with consent (10). The inconsistent findings for
socioeconomic status and health status as predictors of giving consent suggest that these associations could be contextually specific (7-9, 11, 12).

Our findings are subject to certain limitations. First, our trial was powered to detect a 10% difference between groups and thus would not be expected to detect smaller effect sizes as in the present study. Second, as the randomised trial was nested within a cohort, our findings may have limited generalisability to de novo studies that are contacting participants for the first time. In those settings, the effect of requesting unique personal identifiers and provision of incentives may be larger. However, our original sample was randomly drawn from the community and therefore generalisable to other population-based studies.

In conclusion, our trial demonstrated that the request for a unique personal identifier did not substantially reduce consent proportions. Our findings add to the literature on demographic variation in study participation (1), in that there may well be age-based variability for providing sensitive identifying information. Although only one-third of those approached consented to health record linkage, postal invitations could still be a relatively inexpensive, useful initial strategy for cohort studies.
Figure 1. Randomised Allocation of Request for Hong Kong Identity Card number (HKID) and Souvenir Incentive.
Figure 2. Effect of Requesting Hong Kong Identity Card number (HKID) According to Demographic Subgroups. P values were obtained from the likelihood ratio tests of the interaction terms of requesting HKID and the subgroups.
Figure 3. Effect of Providing Souvenir Incentives According to Demographic Subgroups. P values were obtained from the likelihood ratio tests of the interaction terms of souvenir incentives and the subgroups.
Figure 4. Cumulative and Daily Consent Rates. The upper panel shows the temporal associations of receiving completed consent forms with each of the two reminder mailings. The lower panel shows time points where there was a change in daily consent rates as indicated by breakpoints between fitted lines.
## Table 1. Baseline Characteristics of the Study Participants

<table>
<thead>
<tr>
<th></th>
<th>HKID request</th>
<th></th>
<th>No HKID request</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Incentive</td>
<td>No incentive</td>
<td>Incentive</td>
<td>No incentive</td>
</tr>
<tr>
<td><strong>Female sex</strong></td>
<td>No.</td>
<td>%</td>
<td>No.</td>
<td>%</td>
</tr>
<tr>
<td>*</td>
<td>182</td>
<td>62.3</td>
<td>188</td>
<td>61.4</td>
</tr>
<tr>
<td><strong>Age group, years</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18-44</td>
<td>70</td>
<td>24.0</td>
<td>75</td>
<td>24.5</td>
</tr>
<tr>
<td>45-64</td>
<td>131</td>
<td>44.9</td>
<td>135</td>
<td>44.1</td>
</tr>
<tr>
<td>≥ 65</td>
<td>91</td>
<td>31.2</td>
<td>96</td>
<td>31.4</td>
</tr>
<tr>
<td><strong>Educational attainment</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>106</td>
<td>36.3</td>
<td>111</td>
<td>36.3</td>
</tr>
<tr>
<td>Secondary</td>
<td>145</td>
<td>49.7</td>
<td>149</td>
<td>48.7</td>
</tr>
<tr>
<td>Tertiary</td>
<td>41</td>
<td>14.0</td>
<td>46</td>
<td>15.0</td>
</tr>
<tr>
<td><strong>Employment status</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Economically inactive</td>
<td>159</td>
<td>54.6</td>
<td>161</td>
<td>52.8</td>
</tr>
<tr>
<td>Unemployed</td>
<td>7</td>
<td>2.4</td>
<td>4</td>
<td>1.3</td>
</tr>
<tr>
<td>Employed</td>
<td>125</td>
<td>43.0</td>
<td>140</td>
<td>45.9</td>
</tr>
<tr>
<td><strong>Monthly income (HKD)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;10000</td>
<td>111</td>
<td>41.4</td>
<td>115</td>
<td>39.2</td>
</tr>
<tr>
<td>10000-19999</td>
<td>78</td>
<td>29.1</td>
<td>82</td>
<td>28.0</td>
</tr>
<tr>
<td>20000-39999</td>
<td>56</td>
<td>20.9</td>
<td>70</td>
<td>23.9</td>
</tr>
<tr>
<td>≥40000</td>
<td>23</td>
<td>8.6</td>
<td>26</td>
<td>8.9</td>
</tr>
</tbody>
</table>

Abbreviation: HKID, Hong Kong Identity Card number.
## Table 2. Association of Demographics and Health Status with Consent to Health Record Linkage

<table>
<thead>
<tr>
<th>Baseline predictors</th>
<th>N</th>
<th>Consent %</th>
<th>Model 1 OR (95% CI)</th>
<th>Model 2 OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Demographics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female sex†</td>
<td>744</td>
<td>32.3%</td>
<td>0.89 (0.70, 1.14)</td>
<td>-</td>
</tr>
<tr>
<td>Age group, years†</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18-29</td>
<td>94</td>
<td>9.6%</td>
<td>1.00 -</td>
<td>-</td>
</tr>
<tr>
<td>30-44</td>
<td>197</td>
<td>28.9%</td>
<td><strong>3.85</strong> (1.89, 8.68)</td>
<td>-</td>
</tr>
<tr>
<td>45-59</td>
<td>423</td>
<td>35.0%</td>
<td><strong>5.08</strong> (2.62, 11.12)</td>
<td>-</td>
</tr>
<tr>
<td>60-74</td>
<td>307</td>
<td>40.7%</td>
<td><strong>6.49</strong> (3.31, 14.29)</td>
<td>-</td>
</tr>
<tr>
<td>≥75</td>
<td>179</td>
<td>33.5%</td>
<td>4.76 (2.34, 10.76)</td>
<td>-</td>
</tr>
<tr>
<td>Education level</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>436</td>
<td>33.3%</td>
<td>1.00 -</td>
<td>1.00 -</td>
</tr>
<tr>
<td>Secondary</td>
<td>587</td>
<td>34.4%</td>
<td>1.05 (0.81, 1.37)</td>
<td>1.13 (0.83, 1.55)</td>
</tr>
<tr>
<td>Tertiary</td>
<td>177</td>
<td>29.4%</td>
<td>0.83 (0.57, 1.22)</td>
<td>1.52 (0.91, 2.55)</td>
</tr>
<tr>
<td>Employment status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Economically inactive</td>
<td>642</td>
<td>32.4%</td>
<td>1.00 -</td>
<td>1.00 -</td>
</tr>
<tr>
<td>Unemployed</td>
<td>19</td>
<td>31.6%</td>
<td>0.96 (0.33, 2.47)</td>
<td>1.12 (0.37, 3.04)</td>
</tr>
<tr>
<td>Employed</td>
<td>537</td>
<td>34.3%</td>
<td>1.09 (0.85, 1.39)</td>
<td>1.25 (0.90, 1.74)</td>
</tr>
<tr>
<td>Monthly household income (HKD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;10000</td>
<td>477</td>
<td>30.8%</td>
<td>1.00 -</td>
<td>1.00 -</td>
</tr>
<tr>
<td>10000-19999</td>
<td>310</td>
<td>33.2%</td>
<td>1.12 (0.82, 1.52)</td>
<td>1.18 (0.84, 1.65)</td>
</tr>
<tr>
<td>20000-39999</td>
<td>225</td>
<td>37.8%</td>
<td>1.36 (0.98, 1.90)</td>
<td><strong>1.52</strong> (1.05, 2.20)</td>
</tr>
<tr>
<td>≥40000</td>
<td>88</td>
<td>39.8%</td>
<td>1.48 (0.92, 2.36)</td>
<td>1.53 (0.90, 2.58)</td>
</tr>
<tr>
<td><strong>Health status</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical well-being (0-100)</td>
<td>-</td>
<td>-</td>
<td>1.00 (0.98, 1.01)</td>
<td>1.00 (0.98, 1.02)</td>
</tr>
<tr>
<td>Mental well-being (0-100)</td>
<td>-</td>
<td>-</td>
<td>1.00 (0.98, 1.01)</td>
<td>1.00 (0.98, 1.01)</td>
</tr>
<tr>
<td>Diagnosed with chronic metabolic or psychiatric disorders*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>876</td>
<td>32.6%</td>
<td>1.00 -</td>
<td>1.00 -</td>
</tr>
<tr>
<td>Yes</td>
<td>324</td>
<td>34.9%</td>
<td>1.10 (0.84, 1.44)</td>
<td>0.95 (0.70, 1.29)</td>
</tr>
<tr>
<td>Admitted to a hospital in past year</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1114</td>
<td>33.7%</td>
<td>1.00 -</td>
<td>1.00 -</td>
</tr>
<tr>
<td>Yes</td>
<td>86</td>
<td>27.9%</td>
<td>0.76 (0.46, 1.23)</td>
<td>0.77 (0.46, 1.28)</td>
</tr>
</tbody>
</table>

Abbreviation: OR, Odds ratio; CI, Confidence interval.

Model 1: unadjusted
Model 2: adjusted for age, sex, education level, employment status and household income
† We present unadjusted models only for age and sex as other covariates could not be common causes of these exposures and outcomes
* Self-reported doctor-diagnosed chronic metabolic diseases (hypertension, diabetes and hyperlipidaemia) or psychiatric disorders (depression, anxiety disorder or schizophrenia)
ACKNOWLEDGMENTS

The Hong Kong Jockey Club Charities Trust funded the establishment of the original cohort from 2007 to 2014. We thank Dr. Herbert Pang, Dr. Brandford Chan, Ms. Cynthia Yau and Ms Betty Yuan for comments and technical assistance.

Competing interests: None
References


5. Huang N, Shih SF, Chang HY, Chou YJ. Record linkage research and informed consent: who consents? BMC Health Serv Res. 2007;7:18.


