

106 年「高中、高職、五專學生健康行為調查」調查

抽樣設計

一、抽樣母體

本署依據 104 學年度高中、高職、五專（1-3 年級）在校學生（簡稱高中職學生）為母群體，透過教育部網站分別取得高中職學校資料，依學校類型分成四類：1.高中類-全校均為普通科學生的學校、2.高職類-全校均為職業類科學生的學校、3.綜合類-校內普通科、職業科學生均有的學校、4.夜間部-在晚間上課之學生，包含普通科、職業類科學生。依照學校所在鄉鎮市區之人口密度分為四層後，各層採抽出機率與單位大小成比例之等機率抽樣方式，而各層預定抽取之樣本學校數係由事先決定的 60 所樣本學校數按各層學生數佔總學生數之比例進行分配，例如：行政區域為區，學校類型為高中類，則該層預定抽出之樣本學校數的計算方式為： $60 * (58,688 / 829,131) = 4$ (所)。最後，將本署預定樣本學校數及抽樣母體委由美國 CDC 進行抽樣。表 1 為本調查各層預定樣本學校數分配情形。

表 1 106 年高中職學生健康行為調查預定樣本學校數

分層	學校類型	學生數	60 所樣本學校數比例轉換	預定樣本學校數
區	高中類	58,688	4.25	4
	高職類	16,935	1.23	1
	綜合類	50,079	3.62	4
	夜間部	11,372	0.82	1
市	高中類	63,600	4.60	5
	高職類	18,167	1.31	1
	綜合類	77,767	5.63	6
	夜間部	10,713	0.78	1
鄉	高中類	72,594	5.25	5
	高職類	38,394	2.78	3
	綜合類	82,668	5.98	6
	夜間部	14,989	1.08	1
鎮	高中類	59,837	4.33	4
	高職類	105,890	7.66	8
	綜合類	128,670	9.31	9

	夜間部	18,768	1.36	1
合計		829,131		60

二、抽樣設計

本調查抽樣設計係委由美國 CDC 依據全球學生健康行為調查(Global School-based Student Health Survey, GSHS)之取樣方式抽取具全國代表性之樣本，該項調查係採分層二階段集群抽樣設計，依本署提供之教育部 104 學年度高中職教育統計資料為抽樣架構，將全國所有高中職依學校所在鄉鎮市區之人口密度分為四層後，各層採抽出機率與單位大小成比例之等機率抽樣方法(probability proportional to size, PPS)之等機率抽樣方法，先抽出樣本學校，再由樣本學校抽出樣本班級，中選樣本班級之全體學生均為調查對象，約有 6,000 人。

本署依美國 CDC 抽中之樣本學校及其樣本班級代號，於年底函文各縣市衛生局就中選樣本學校 104 學年度各年級實際班級數及班級命名方式進行瞭解後，將樣本學校之所有班級依一、二、三年級(如表 2)，並依序給予班級代號，再與第二階段中選之樣本班級代號核對，即可得知實際中選樣本班級名稱。

表 2 106 年高中職學生健康行為調查抽選樣本班級範例

學校	班級數			樣本 班級數	中選班級 名稱	中選班級 代號
	一	二	三			
○○高中	1 班	1 班	1 班	3	○○高中一年 3 班	3
	2 班	2 班	2 班		○○高中二年 2 班	10
	3 班	3 班	3 班		○○高中三年 4 班	19
	4 班	4 班	4 班			
	5 班	5 班	5 班			
	6 班	6 班	6 班			
	7 班	7 班				
	8 班					

三、樣本加權公式及母體推估公式

(一) 樣本加權公式說明如下：

$$W = W_1 * W_2 * f_1 * f_2 * f_3$$

W_1 =學校抽出率的倒數

W_2 =每所學校的班級抽出率的倒數

f_1 =學校層級：依學校註冊人數分為大中小學校，以進行學校完訪率的校正

f_2 =學生層級：依每班學生完訪率進行校正

f_3 =依性別、年級進行事後分層的校正

(二) 母體推估公式說明如下：

$$1. \hat{P} = \sum_{i=1}^{16} W_i \hat{P}_i$$

\hat{P} ：全國特徵估計值

\hat{P}_i ：第 i 層之特徵值估計

$W_i = W_1 \times W_2 \times f_1 \times f_2 \times f_3 \times f_4$ ：第 i 層之權數

$$2. \text{Var}(\hat{P}) = \sum_{i=1}^{16} (W_i)^2 \text{Var}(\hat{P}_i), \text{ 其中 } \text{Var}(\hat{P}_i) = \frac{P_i(1-P_i)}{n_i} \times \frac{N_i - n_i}{N_i - 1}$$

$\text{Var}(\hat{P})$ ：全國特徵估計值的變異數

$\text{Var}(\hat{P}_i)$ ：第 i 層之特徵值變異數的估計

P_i ：第 i 層之全國特徵值

N_i ：第 i 層之母體人數

n_i ：第 i 層之樣本數