

馬鈴薯病毒病感染率之群體測試

蔣國司¹ 鍾文全^{2,3} 林上湖² 賴信宏¹ 王淑芬¹

¹ 台中市 國立中興大學農藝學系

² 台中縣 行政院農業委員會種苗改良繁殖場繁殖技術課

³ 聯絡作者，電子郵件：wcchung@tss.gov.tw；傳真：+886-4-2582-5818

接受日期：中華民國 97 年 7 月 25 日

摘要

蔣國司、鍾文全、林上湖、賴信宏、王淑芬. 2008. 馬鈴薯病毒病感染率之群體測試. 植病會刊 17: 321-326.

馬鈴薯是世界主要四大糧食作物之一，由於台灣氣候型態高溫多濕，馬鈴薯栽培期間易遭受病毒病害之危害，大大影響馬鈴薯之產業發展。在田間栽培期間，常使用群體測試來估計馬鈴薯病毒病發生率，但在每個群集中，應使用幾株為一檢測單位常是最終影響成本最重要之關鍵。本研究收集行政院農業委員會種苗改良繁殖場所栽種共5年(2004~2008) 8個馬鈴薯田間試驗資料，並應用廣義線性模式來估計單株馬鈴薯病毒病感染率，進而以均方誤差為標準，來計算每群集最佳之植株數，結果顯示可將原來每群集 20 株檢測植株擴增至 40 株，這不僅能降低成本且不影響檢測之精確度，可達到最佳化之目的。

關鍵詞：群體測試、感染率、均方誤差、廣義線性模式

馬鈴薯 (*Solanum tuberosum* L.) 係屬茄科植物，原產於中南美洲地區，是世界主要四大糧食作物之一。台灣馬鈴薯的產地主要集中於台中縣、雲林縣與嘉義縣等三縣，其中以雲林縣斗南地區栽培面積最大，其次為台中縣豐原與嘉義縣溪口地區。由於台灣地處熱帶與亞熱帶地區，高溫多濕，馬鈴薯於栽培期間易遭受許多病蟲害的危害，其中以病毒病害危害較嚴重，因其可藉由無性繁殖體傳播，進而快速且廣泛的影響整個馬鈴薯之產業發展，因而受到產業界、官方機構與研究機關極度重視。馬鈴薯之病毒病害種類多達 25 種，而大多數生產馬鈴薯的國家所遭遇的主要病毒病害，以馬鈴薯A病毒 (*Potato virus A*, PVA)、馬鈴薯 S 病毒 (*Potato virus S*, PVS)、馬鈴薯 X 病毒 (*Potato virus X*, PVX)、馬鈴薯 Y 病毒 (*Potato virus Y*, PVY) 及馬鈴薯捲葉病毒 (*Potato leaf roll virus*, PLRV) 為主⁽¹⁾，因此台灣馬鈴薯健康種薯生產之標的檢測對象亦是針對此五類病毒進行。行政院農業委員會為獲得優良的健康馬鈴薯種薯，持續供應農民生產，特委由種苗改良繁殖場設立健康馬鈴薯種薯生產三級制，即從組織培養、基本種至原原種階段均由種苗改良繁殖場生產，

至於原種與採種兩階段則分別由各地區農會或合作社等單位進行量產的工作。

在過去植物病理的文獻上^(3, 4, 5, 6, 8, 9, 10)，有探討當田間植株感染率 (infected rate) 不大且檢驗之成本所費不貲時，可採取群體測試 (group-testing) 來代替單一植株 (individual unit) 之檢測，其最終結論為群體測試通常比單一測試有效、方便且準確。現今在種苗改良繁殖場是採用 20 株馬鈴薯植株當作一個群集 (group) 檢測單位，但往往造成成本過高，因此是否可讓每個群集有較大數目的植株數，且不影響檢測之精準度，將是本文最主要探討的重點。

每植株所感染病毒之機率定義為，在每一群集中至少有一植株罹病之機率為 $p_{\text{group}} = 1 - (1-p)^k$ ， k 代表每個群集之株數，上述式子整理後可得 $p = 1 - (1-p_{\text{group}})^{\frac{1}{k}}$ ，而 p_{group} 之估計值 $p_{\text{group}} = \frac{\sum_{i=1}^N Y_i}{N}$ ， N 代表群集數，以 2004 年例， N 等於 180；若 $Y_i = 0$ 代表這群集檢測出每植株皆為健康，但若 $Y_i = 1$ 則代表這群集至少有一棵植株罹病⁽⁷⁾。

通常每棵植株之感染率 p 是較每個群集之感染率

p_{group} 有生態上之意義，舉例來說，假使我們想比較兩抗病品種，在相同的條件下每植株之感染率為 $p_1 = 0.1$ 與 $p_2 = 0.2$ ，我們可說感病性後者之品種為前者的兩倍；但假使以 5 株為一單位量測 ($k=5$)，前者罹病率為 $1-(1-0.1)^5 = 0.41$ ，後者為 $1-(1-0.2)^5 = 0.67$ ，故感病性後者品種為前者 1.63 倍；又若 $k=10$ ，則感病性後者品種為前者 1.37 倍，因此，當使用不同之 k 值，結果將隨之改變，一般研究目的皆著眼於估計每植株之感染率 p 。

因此本研究首先將描述田間馬鈴薯病毒病的檢測過程與所進行之 direct-ELISA 流程，並進而闡述如何使用群體檢測，來估計每株之感染率及計算每群集之最佳株數，並以此探討成本效應之分析，最後說明及建議群體測試之適用範圍。

本研究共在 2004、2005、2006、2007 與 2008 年 3 月 5 日進行馬鈴薯田間病毒之檢測 (表一)，除了 2005、2006 與 2007 年第二次栽種品種為種苗 2 號，其餘皆為克尼伯，因兩者皆為感病品種，故本研究將其一併考慮分析；所有栽種模式皆為田間栽培，雖群集數有所不同，但皆大於 100 (除了 2007 年第二次的群集數為 99)，其每個群集之株數皆固定，即 20 株/群集，在檢測過程中針對 20 株之每株馬鈴薯隨機抽取一葉片，將此 20 個葉片混合研磨，進行直接酵素聯結抗體免疫法 (direct-ELISA)，檢測上述之五種病毒是否為陽性反應。因本研究是在探討每個群集所包含植株之數量對其最終精準度之影響，故將其檢測結果視為健康或罹病，並不特別針對每類病毒加以探討。

將 PVA、PVS、PVX、PVY 及 PLRV 等五種馬鈴薯病毒病之抗體，分別以敷膜溶液 (coating buffer: Na_2HCO_3 2.93 g, Na_2CO_3 1.59 g, 去離子水 1000ml pH = 9.6) 稀釋成 1000 倍，然後以八爪微量吸注器吸取 100 μl 至微量盤 (ELISA plate) 中，再將微量盤至入保

濕盒中，於 37 °C 反應 2 小時後，取出以 1X PBST (NaCl 8.0 g, Na_2HPO_4 1.15 g, KH_2PO_4 0.2 g, KCl 0.2 g, Tween-20 0.5 ml, NaN_3 0.2 g, 去離子水 1000 ml, pH = 7.4) 清洗 3 次，每次 5 分鐘，再將微量盤拍打至厚紗布上用乾。將測試樣品稱重 0.5 g，並將植物組織與萃取液 (Polyvinylpyrrolidone (PVP) 20 g, Powdered egg (chicken) albumin, Grade II 2 g, Na_2SO_3 1.3 g, Tween-20 20 ml, NaN_3 0.2 g, PBST 1000 ml, pH = 7.4)，依 1:10 (W/V) 之比例置於研磨袋中研磨，再以微量吸注器吸取組織研磨液 100 μl 分別注入 96 格微量盤中，每樣品 2 重複，並以健康馬鈴薯組織液 (負對照) 及含病毒馬鈴薯組織液 (正對照) 當作對照組，然後將微量盤放入保濕盒中，經 4 °C 過夜後，以洗滌液 1X PBST 清洗微量盤 5 次，每次 5 分鐘，同樣用乾。隨後將 enzyme-conjugated antibody 以結合緩衝液 (BSA 2.0 g, PVP-40 20 g, Na azide 0.2 g, PBST 1000 ml, pH = 7.4) 稀釋成 1000 倍，再以八爪微量吸注器吸取 100 μl 至微量盤中，並將微量盤至入保濕盒中，於 37 °C 反應 4 小時後，取出以 1X PBST 清洗 3 次，每次 5 分鐘，同樣用乾。最後將 PNP (p-nitrophenyl phosphate disodium) 與基質緩衝液 (Diethanolamine 97.0 ml, Na azide 0.2 g, $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$ 0.1 g, 去離子水 1000 ml, pH = 9.8) 依 1:1 (mg/ml) 的比例配製成鹼性磷酸酵素基質試液後，以八爪微量吸注器吸取 100 μl 加入至微量盤中，再放入保濕盒內，於 37 °C 暗室靜置 30 分鐘 ~ 1 小時後取出，以 ELISA 讀值光譜分析儀分析 O.D. 405nm/495nm 之吸收值。

自冷藏庫挑取馬鈴薯每粒種薯重約 25 公克的克尼伯品種，整齊排列置於塑膠籃內 (長×寬×高，48×28×20 cm)，並於室內陰涼處進行日光燈照光育芽，經 6 星期後，將已發芽的馬鈴薯種薯，分別種植於行政院

表一、2004 至 2008 年田間馬鈴薯 5 種病毒之檢驗調查與群體感染率估計值

Table 1. Results of 2004-2008 surveys and testing of potatoes for the five viruses showing the number of samples infected and estimate of groups infected

Time (section)	Number of groups (N)	Numbers of samples infected							Estimate of groups infected ² (%) (p_{group})
		PVA	PVS	PVX	PVY	PLRV	Healthy	Infection ¹	
2004 (1)	180	0	0	0	0	0	180	0	0
2005 (1)	291	1	5	0	10	2	277	14	0.048
(2)	180	0	0	0	0	0	180	0	0
2006 (1)	152	0	4	0	0	1	148	4	0.026
(2)	150	0	0	0	0	0	150	0	0
2007 (1)	313	0	13	0	13	0	290	23	0.073
(2)	99	0	4	0	21	21	68	31	0.313
2008 (1)	513	1	2	0	4	3	503	10	0.019

¹ The numbers of samples were infected by at least one virus.

² Twenty plants were used as a group.

農委會種苗改良繁殖場的網室田間，共計 11 畦，每畦（長×寬，45×1.2 m）採雙行栽植，並以三角形方式種植種薯，薯距 45 cm 共計種植 200 顆馬鈴薯種薯。馬鈴薯種薯種植後當日覆土 5-7 cm，然後水分管理採用 3 孔塑膠噴水帶進行噴灌，於生長中後期約 5 至 7 天灌溉 1 次，收穫前 10 天即停止灌溉，並施以一般的病害管理及施肥，待馬鈴薯種薯採收時前 1-2 星期，採收馬鈴薯植株以 direct-ELISA 進行檢測。

Swallow 曾詳述群體測試在植物病理調查的使用⁽⁹⁾，而導出以下之式子， $E(\hat{p})$ 代表期望值， $\text{Var}(\hat{p})$ 代表變異數及 $\text{MSE}(\hat{p})$ (mean square error) 代表均方誤差：

$$\text{Bias}(\hat{p}) = E(\hat{p}) - p \quad (1)$$

$$\text{Where } E(\hat{p}) = 1 - \sum_{i=0}^N \left(\frac{i}{N}\right)^k \binom{N}{i} [(1-p)^k]^i [1-(1-p)^k]^{N-i}$$

$$\text{Var}(\hat{p}) = E[\hat{p} - E(\hat{p})]^2 \quad (2)$$

$$= \sum_{i=0}^N \left(\frac{i}{N}\right)^k \binom{N}{i} [(1-p)^k]^i [1-(1-p)^k]^{N-i} - [1 - E(\hat{p})]^2$$

$$\text{MSE}(\hat{p}) = E(\hat{p} - p)^2 = \text{Var}(\hat{p}) + [\text{Bias}(\hat{p})]^2 \quad (3)$$

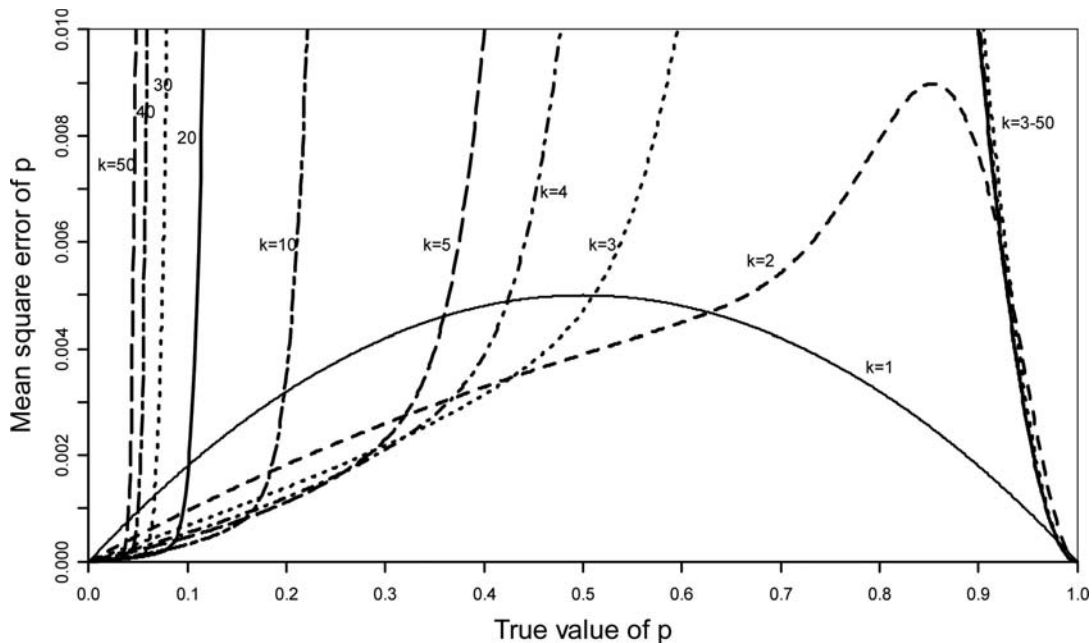
均方誤差 (MSE) 常代表估計值之準確性 (accuracy) 與精確度 (precision)，前者代表誤差 (bias)，後者代表變異數 (variance)，僅有當誤差小且變異數小時，均方誤差才變小，因此在統計上它常被當成一估計值優劣

之量測標準。從式 (3) 我們得知估計值之均方誤差與群集數 (N)、每一群集之株數 (k) 及真正每株植株之感染率 (p) 有關，給予設定 $N=50, k=1, 2, 3, 4, 5, 10, 20, 30, 40, 50$ ，可見其 p 與 $\text{MSE}(p)$ 的關係 (圖一)。

本研究著手計算在於不同的群集數 (N) 與每植株之感染率 (p) 的條件下，試求在每一群集中最佳的株數 (k)，亦即 MSE 為最小時之 k 值。在此我們是以 R 程式語言來完成此一研究 (程式名稱：group testing)，只要輸入 N 與 p 值，即可得最佳之 k 值 (從 1 至 50)，因超過 50 在田間調查時並不常見，亦恐會違反其假設前提，在文後討論時將再詳述。

近期發展出的 R 程式語言為一免費的統計分析軟體，包含許多數學、統計與圖形的函數，可據此書寫有興趣主題的程式，其最大優點為取得軟體的便利性，並具有一般統計軟體 (例如：SAS) 分析數據的強大功能，因此近幾年在學術領域上的運用相當普遍，在美國植物病理學會 (The American Phytopathological Society) 教育中心 (education center) 的網站上也有介紹植物流行病學與生態學的課程，全部內容皆以 R 程式來加以說明與量化，可見 R 程式語言已普及至植物保護的領域。

本群體測試中，最佳之設計必須事先知道馬鈴薯之感染率 (p)，但這正是我們所要估計的值，Thompson⁽¹²⁾ 與 Swallow^(9, 10) 曾提出以 p 值大概之上界



圖一、每群集有 k 植株且共有 50 個群集數情況下，單一植株感染率對於相對應之均方誤差關係圖， $k=1$ 代表單株檢測而 $k>2$ 代表使用群體測試檢測。

Fig. 1. Mean square error of the true infection rate (p) versus p for a group of k plants, $k=1, 2, 3, 4, 5, 10, 20, 30, 40, 50$ with total number of groups, $N=50$. $k=1$ represents that all plants are tested individually; $k>2$ represents that all plants are tested by group-testing.

來代替 p 值，即可求得 k 值，雖說所求得 k 值過於保守，將小於最佳之 k 值，但它表現能力遠遠大於單一植株之檢測。因此本文首先由之前資料去估計 p 值，並賦予相當保守之 p 值上界，再代入上述 group-testing 之 R 程式，求得最佳之 k 值，並與現今所採用之 k 值 ($k = 20$) 作一比較再討論反映至成本效應問題上。

從(圖一)，可知當 p 值為小時， $k > 1$ 比 $k = 1$ 之 MSE 還小，有時更相差數倍，甚至 k 值越大，在某些情形 MSE 反而小於 k 值較小者，這和我們直覺認知不盡相同，且 p 值越小，最佳化之 k 值越大，從(圖一)得知當 p 介於 0 與 0.62 時，群體測試 ($k > 1$) 較優於單株測試 ($k = 1$)。

本研究使用 Farrington⁽²⁾ 之概念，將群體測試視為 Complementary Log-Log (CLL) 之形式，即 $p_{\text{group}} = 1 - (1 - p)^k$ ，可導至 CLL (p_{group}) = $\ln(k) + \text{CLL}(p)$ ，這裡 CLL (●) 定義為 $\ln[-\ln(1 - \bullet)]$ ，因此可結合 5 年共 8 筆田間資料來估計每植株之感染率 (p)。從(表一)我們可觀察到 p_{group} 變動不小(從 0 至 0.3 左右)，因此必須注意是否有過度離勢 (overdispersion) 現象，即觀測之變動遠遠大於理論之變動程度(在此為二項分佈之變量)，此處用 R 程式來分析，結果所估計 CLL (p) 為 -5.79，其標準差 (standard error) 為 0.64，經過轉換 p 為 0.003，及標準差為 0.002，95% 信賴區間為 (0, 0.007)，其結果的 deviance 為 5.27 (自由度為 7)，故 p value 為 0.63，亦即配適相當不錯，此處考慮到過度離勢現象^(2, 13)，若無考慮過度離勢現象，則配適結果不佳。

雖然從 5 年期間的 8 筆評估試驗結果發現，每棵馬鈴薯植株病毒感染率為 0.003，但根據 Swallow^(9, 10) 建議可採用 p 值之上界而得一較小 k 值 (每群有 k 株)，因在(表一)資料中群體感染率介於 0 至 0.3 間，藉著 $p_{\text{group}} = 1 - (1 - p)^k$ 換算成單一植株感染率為 0 至 0.02 之間，接續我們使用撰寫的程式，當使用 p 值為 0.02 時，最佳 k 值為 50；而 p 值為 0.03 時，最佳 k 值為 47； p 值為 0.04 時，最佳 k 值為 35，這幾乎是原來田間檢測所採用之 k 值的兩倍，因此從這研究結果得知，當使用 $p = 0.03$ 與 0.04 為真正 p 值之上界時，則 k 依序為 47 與 35，MSE 為其最小，為最佳之每群株數；又在群體測試中， N 越大則 k 值越大，例如： $N = 100$ 和 200 時，若 p 皆等於 0.03 代入，則 k 值依序為 47 與 50，在(表一)中 N 大部份皆遠大於 100，故在此我們相當保守使用 $p = 0.03$ 或 0.04 來估其每株之感染率，其結果相當具有可信度。

若採用本研究之結果，每群之株數大約為 40，可減少原來成本接近一半 (因原來 $k = 20$)，如此一來，便大大節省農民之成本，因此藉著群體測試，使用者可

在最經濟及效率之條件下，選擇最佳之群體測試來完成田間管理之工作。

植株之檢測以群集為單位非以單株為單位，稱之為群體測試。過去之文獻已證明^(9, 10, 11)，群體測試在某些情況下，表現總是優於單一檢測，例如：在(圖一)真實感染率介於 0 至 0.62 時的情形，所用的判斷標準為 MSE，因 MSE 是估計誤差平方與變異數之總和。

在農業委員會種苗改良繁殖場為了要供應各區農會或合作社馬鈴薯之健康原原種，故必須對田間所栽培之植株加以檢測是否帶有病毒，在 2004 ~ 2008 年間共 8 個田間栽培皆以每 20 株為一單位進行之前所描述 5 種常見病毒檢測 (大約新台幣 190 元/群集單位)，因此造成甚大的成本負擔，本研究即考慮成本效應，希望能將成本降低，因每群集大於 20 株為單位，其成本相同於 20 株為一單位，但 k 值是否能兼顧到成本又不喪失其精確度呢？群體測試正可解決此一問題。

由之前種苗改良繁殖場所收集之 5 年田間資料可發覺群體病毒感染率並不太高 (至多 0.3)，故使用群體測試應用於此一問題相當適切，因群體測試 MSE 在某些情況甚至小於單株測試之 MSE；但在估計每群體最佳株數時，需事先知道單株之感染率，這是未知數正是所要估計之值，在此使用廣義線性模式 (Generalized linear model) 有二項式誤差 (binomial error) 與 Complementary Log-Log (CLL) 之連接函數 (link function) 來估計 p ， $\ln(k)$ 之係數設定為 1，則截距項 (intercept) 即為 CLL (p)，經過轉換每株之感染率 (p) 即為所求，並可同時得到其標準差與信賴區間；因不同的田間試驗結果，其群體感染率有不小的變化介於 0 至 0.3 之間，在此遂採用準概度法^(2, 13) (quasi-likelihood methodology) 來進行分析，所得配適結果相當不錯，進而求得單株感染率 (p) 為 0.003。不少探討群體測試之文獻^(9, 10, 12) 曾建議在推估每群之株數時要保守估計 p 值，因此我們使用撰寫的程式，當採用 $p = 0.02$ 時， k 為 50；若 $p = 0.03$ ， k 為 47；若 $p = 0.04$ ， k 為 35；此程式是用 R 程式加以撰寫，只要輸入群集數 (N) 與單株感染率 (p)，即可求得每群集最佳之株數 (k)，其最佳化之標準為 MSE。

一般而言，在田間試驗 k 值甚少採用超過 50，因為若採用甚大的 k 值，恐造成檢測結果皆為陽性，如此將大為增加對於單株感染率估計之誤差及 MSE，所得結果不甚有用^(3, 7, 9, 10, 11)。因此，根據上述的理由我們建議 k 大約為 40 為最佳之設計，即每群有 40 株馬鈴薯植株為單位進行 direct-ELISA 檢測，因以 20 或 40 株檢測為單位其成本皆相同，且 40 株為單位其效率不輸 20 株 (因 MSE 較小)，故檢測成本可減少一半，大幅

降低成本所造成的負擔，並可將此減少之成本嘉惠給生產單位來增進農民之福祉。

在說明群體及單一感染時，我們使用 $p_{\text{group}} = 1 - (1 - p)^k$ 代表彼此之關係，這是假設每植株彼此感病的機率是固定且獨立的，故可用二項分佈來描述，但田間的生態並非每植株間都符合獨立的性質，以馬鈴薯病毒病為例，媒介昆蟲為其傳染病毒首要之惡，但媒介昆蟲常有群聚性 (aggregated)，二項分佈並不適合此類田間現象之描述，貝他二項分佈 (beta-binomial distribution) 能描繪其群聚性，應是較佳之選擇，國外文獻曾以大桔蚜 (*Toxoptera citricida*) 感染柑橘萎縮病 (Citrus Tristeza) 為例來加以闡述此現象⁽⁹⁾，這也正是我們所要進行的下一步研究。

群體測試在本文應用於馬鈴薯病毒病檢測之成本分析，對於這類方法我們著眼於單株感染率之估計與每群之最佳株數，其實群體測試研究也有另一方向，即尋找那株為染病之植株，在植物病理研究上常碰到此類的情形，這是重新測試 (retest) 的問題⁽³⁾，但並非為本文所關心的主題。

群體測試可證明單株檢測的精確度並非永遠是最大的，當單株感染率甚小時尤其明顯，此方法應不限於本文所描繪的應用，各類病蟲害的調查，其感染率甚小且成本花費不少，群體測試當可發揮其最大效能。

引用文獻 (LITERATURE CITED)

1. Loebenstein, G. 2007. Potato virus diseases, their diagnosis and preparation of virus-tested seed potatoes by rapid propagation. Acta Hort. (ISHS) 729:437-440.
2. Farrington, C. P. 1992. Estimating prevalence by group testing using generalized linear models. Stat. Med. 11:1591-1597.
3. Hepworth, G. 1996. Exact confidence intervals for proportions estimated by group testing. Biometrics 52:1134-1146.
4. Hughes, G., and Gottwald, T. R. 1998. Survey methods for assessment of citrus tristeza virus incidence. Phytopathology 88:715-723.
5. Hughes, G., and Gottwald, T. R. 1999. Survey methods for assessment of citrus triteza virus incidence when *Toxoptera citricida* is the predominant vector. Phytopathology 89:487-494.
6. Hughes, G., Gottward, T. R., and Levy, L. 2002. The use of hierarchical sampling in the surveillance program for *Plum pox* virus incidence in the United States. Plant Dis. 86:259-263.
7. Madden, L. V., Hughes, G., and van den Bosch, F. 2007. The study of plant disease epidemics. The American Phytopathological Society, APS Press, St. Paul, Minnesota. 421 pp.
8. Rodoni, B. C., Hepworth, G., Richardson, C., and Moran, J. R. 1994. The use of a sequential batch testing procedure and ELISA to determine the incidence of five viruses in Victorian cut-flower Sim carnations. Aust. J. Agric. Res. 45:223-230.
9. Swallow, W. H. 1985. Group testing for estimating infection rates and probabilities of disease transmission. Phytopathology 75:882-889.
10. Swallow, W. H. 1987. Relative mean squared error and cost considerations in choosing group size for group testing to estimate infection rates and probabilities of disease transmission. Phytopathology 77:1376-1381.
11. Tebbs, J. M., and Bilder, C. R. 2004. Confidence interval procedures for the probability of disease transmission in multiple-vector-transfer designs. J. Agric. Biol. Environ. Stat. 9:75-90.
12. Thompson, K. H. 1962. Estimation of the proportion of vectors in a natural population of insects. Biometrics 18:568-578.
13. Williams, D. A. 1982. Extra-binomial variation in logistic linear models. Appl. Stat. 31:144-148.

ABSTRACT

Chiang, K. S.¹, Chung, W. C.^{2,3}, Lin, S. H.², Lai, H. H.¹, and Wang, S. F.¹ 2008. Group-testing design for the infection rates of potatoes viruses. *Plant Pathol. Bull.* 17: 321-326. (¹ Department of Agronomy, National Chung Hsing University; ² Propagation Technology Section, Taiwan Seed Improvement and Propagation station; ³ Corresponding author, Email: wcchung@tss.gov.tw; Fax: +886-4-2582-5818)

Potato (*Solanum tuberosum* L.) is one of the four main foods in the world. Because of the climate of Taiwan, the cultivation of potato suffered from the virus disease easily, further influencing the development of potato industry. Group-testing was often utilized to estimate the incidence rate of potatoes viruses in the field. It is the principal issue that how many group sizes should be regarded as a testing unit for group-testing because it is usually related to the cost investigated. We collected eight field data sets of potatoes farms in five years (2004 ~ 2008) from Taiwan Seed Improvement and Propagation Station. Besides, generalized linear models were employed to estimate the proportion of the individual plant infected. By mean square error (MSE) as criteria, optimal group sizes were changed from 20 to 40 plants per group so that the cost was reduced and the testing process was improved efficiently. Not only can the group-testing design advance the accuracy of testing, but it can reach to the purpose of optimization.

Key words: group-testing, infection rate, mean square error, generalized linear model