

植物病害量化順序尺度之優化與統計分析方法之進展

蔣國司¹、洪爭坊^{2,*}

¹ 國立中興大學農藝學系

² 國立中興大學植物病理學系

* 聯絡作者，E-mail: cfhong@nchu.edu.tw

摘要

蔣國司、洪爭坊。2025。植物病害量化順序尺度之優化與統計分析方法之進展。植物醫學67(2): 69-78。

植物病害的量化評估，對於作物病害管理、抗病育種、以及病害預警系統等資料應用與相關領域的研究至關重要。常見的病害量化評估方法包括評估病害盛行率、罹病率與罹病度。其中，過去針對罹病度的定義為病徵佔植物組織或器官的百分比，因此常評估病徵的最近似百分比(nearest percent estimate; NPE)作為罹病度的估計值。然而，並非所有評估者都能快速準確的評估NPE，因此許多研究提出可以使用量化順序尺度進行罹病度的評估。其中，Horsfall-Barratt scale 即為一例，該尺度在過去已被廣泛使用在不同病害的研究上。然而，由於其非線性級距的設計，許多研究陸續提出此類尺度的設計結構容易引入誤差，進而影響後續的統計檢定力。本文回顧過去的研究，除了探討如何優化量化順序尺度的設計結構，以提升罹病度評估的準確度，同時介紹對於量化順序尺度資料統計分析方法的近期研究成果，期能使植物保護相關研究人員對此類資料的收集與分析能有進一步的認識。

關鍵詞：植物病害量化、罹病度、量化順序尺度、比例勝算模型、存活分析

前言

植物病害的量化資料，在評估作物病害管理效果、植物病害的風險與可能帶來的經濟損失、作物抗病育種、建立病害預警模型，以及許多後續的應用中，都扮演關鍵的角色。傳統上，病害量化資料常包括評估病害的盛行率(disease prevalence)、罹病率(disease incidence)、與罹病度(disease

severity)^(5, 7, 29)。其中，盛行率主要在描述較大的空間範圍內，在特定數量或比例的地理或行政區單位中已經發生某個病害。例如：雲嘉南地區在5月份的稻熱病鄉鎮盛行率為20%，亦即雲嘉南地區，有20%左右的鄉鎮在5月份時已經發現稻熱病。罹病率，主要描述在總共調查的作物樣本中，有多少比例的樣本已經受到特定病害危害。例如：調查80顆番石榴果實中，有43顆果實有病害，因此該番石榴果實樣本的罹病率為 $43/80=53.8\%$ 。前述兩種量化評估的方法較為簡單，且直接評估地區或樣本中「有、無」病害發生，因此對於相同或不同評估者而言，評估的結果較為一致。然而，前述的量化資訊，僅說明調查地區或樣本中「有、無」病害，並不包含各病害「發生多嚴重」的訊息。因此，就會以罹病度來量化或描述病害的嚴重程度。

罹病度常被定義為：病徵在調查的組織、器官或植株樣本上所佔面積的比例。例如：柑橘潰瘍病的病徵佔柑橘葉片面積約12%，因此罹病度為12%，或記作0.12。此種評估方式稱為最近似百分比評估值(nearest percent estimate; NPE)，至今仍主要仰賴人力進行判斷。然而，實務上由於人力資源與時間的限制，此種仰賴目視估算的方式，常會因評估者本身或評估者之間的差異，以及眾多的因素，而導致較大的評估差異。為了克服前述的限制，許多學者提出各式罹病度等級尺度(disease severity scale)以輔助將不同比例的罹病面積，區分成不同的等級，主要目的除了讓使用者能更快評估病害的嚴重程度，也儘量降低評估者內或評估者間在評估NPE時的差異。在過去提出的各種罹病度等級尺度中，最具代表性者屬 Horsfall-Barratt Scale (HB scale)(表一)，此種方法被歸類為量化順序尺度(quantitative ordinal scale)^(5, 14)，本文將著重於這一類的方法深入探討。此外，亦可直接評估病斑的數量、單位面積上的病斑密度、病斑直徑、長度等代表罹病度，然而這些方法各有其優缺點，且不一定適用於所有的病害，在此不再深入介紹。除此之

表一、Horsfall-Barrat 罹病度等級尺度所對應的病害面積百分比範圍、級距大小與中位數^(3, 29)

TABLE 1. Horsfall-Barrat Scale and corresponding disease severity range, interval size, and midpoint^(3, 29)

HB Scale	Disease severity range	Interval size	Midpoint
1	0	0	0
2	0 ⁺ -3%	3	1.5
3	3 ⁺ - 6%	3	4.5
4	6 ⁺ - 12%	6	9.0
5	12 ⁺ - 25%	13	18.5
6	25 ⁺ - 50%	25	37.5
7	50 ⁺ - 75%	25	62.5
8	75 ⁺ - 88%	13	81.5
9	88 ⁺ - 94%	6	91.0
10	94 ⁺ - 97%	3	95.5
11	97 ⁺ - 100%	3	98.5
12	100%	0	100

外，有些病害較不容易直接評估病徵佔組織或器官的比例或面積大小，例如：青枯病、萎凋病、病毒病、線蟲造成的根瘤…等(圖一)，此類病害則較適合以質化順序尺度(qualitative ordinal scale)進行描述與評估⁽⁵⁾，在此亦不多作說明。

由前述介紹可知，作物病害的種類繁多且差異甚大，因此在進行病害調查時，應先確認病害量化資料後續的用途，以選擇較符合目的的調查方法。若進行罹病度的評估時，除了選用較合適的方法外，也應留意後續的資料分析步驟，以獲得正確的結論。本文主要介紹並回顧量化順序尺度理論及其在實務上的應用相關研究。這些研究透過模擬與田間數據，驗證不同罹病度等級尺度的結構對估算準確性與統計檢定力的影響，並進一步探討適用於罹病度等級尺度的設計與後續資料統計分析的方法。

罹病度等級尺度的發展

植物病害的嚴重度評估自19世紀末期開始即受到重視，尤其在許多研究中，如何快速且準確地進行病徵的量化，已成為探討的重點之一。一般認為較理想的方式為直接評估病斑在植物組織或器官上所佔面積的比例，但由於操作不容易且評估者之間的主觀差異大，且評估時容易受到多種因素干擾，因



器官：包括但不限於整棵植株、根系、枝條、葉片、花、果實、塊根、塊莖…等。

病徵：包括但不限於萎凋、黃化、嵌紋、斑駁、矮化、腫大、變形…等，能以文字描述但不易測量的症狀。

圖一、不同的病害種類以及可能適用的罹病度評估方式。合適的罹病度評估方法，應視病害種類與研究目的等因素決定。

Fig. 1. Different types of diseases and the methods for assessing disease severity. The method selected for disease severity assessment should be based on the patho-system and the objectives of the study.

此後續許多研究中，提出採用順序尺度(ordinal scales)作為輔助評估罹病度的替代方案，認為可以較快速的協助評估者量化病害並降低誤差^(4, 16, 24)。常見將罹病面積百分比轉換為等級所使用的罹病度等級尺度，例如：0級=0%；1級=0⁺~10%；2級=10⁺~25%；3級=25⁺~50%；4級=50⁺%⁽²⁵⁾。依照不同病害種類，研究人員亦常自行定義不同等級所代表的病徵面積百分比^(19, 20, 26, 28, 39)。其中，Horsfall-Barratt Scale (HB scale)是較為廣泛使用的罹病度等級尺度。該方法由 Horsfall 與 Barratt 於 1945 年提出，稱係依據心理物理學(psychophysics)的Weber-Fechner law所設計，強調目視對病徵的辨識能力，在病斑面積小於50%時，能較有效的辨識病斑大小，而病斑面積大於50%時，則較能辨識健康部位的大小；並且罹病度等級尺度的級距設計，應該呈對數的變化而不是線性的變化^(4, 17, 24, 30)。

罹病度等級尺度在使用上雖然方便，卻可能因為各等級對應的罹病面積的級距非線性分佈，導致在直接使用罹病度等級分析資料時導入誤差，進而影響後續假設檢定與研究的結論，前述的不同罹病度等級尺度亦是如此。以HB scale為例，Redman等(1968)提出改採各等級所對應的級距中位數(midpoint)，作為各級代表值，例如：以37.5%代表 25 – 50%，計算平均值之後再進行分析^(4, 35)。然而，後續的多個研究結果則認為使用HB scale時，評估者在評估該範圍的面積百分比時，評估值與正確值的關係並非對數關係，而是呈線性關係，且因為接近50%的罹病度等級尺度的級距較寬，因此使用HB scale無法獲得比NPE較正確且精準的結論^(17, 30, 31)，且在某些情形下較容易產生第二型錯誤(Type II error)，亦即在虛無假設(H₀)為假的情況下，會因無法拒絕虛無假設而降低統計的檢定力(power) (無法區別處理間的差異)^(2, 3, 17, 30, 31)。因此，近年來逐漸有學者利用田間資料或模擬分析，驗證不同尺度轉換對罹病度評估結果的影響，並針對罹病等級尺度的設計結構與後續分析方式提出優化的作法^(4, 9, 30, 31)。總而言之，HB scale 與其衍伸的罹病度等級尺度雖然在植物病理學界已被廣泛應用，但該尺度的結構與後續資料的分析方法，仍有進一步優化的空間。

如何優化罹病度等級尺度的結構？

在進行植物病害評估後，研究者往往需要比較不同資材、品種、栽培方式或環境等因子對於罹病度的影響，進而進一步做決策，即假設檢定(hypothesis testing)。然而，若利用不等距的罹病度等級尺度評估病害，並直接以罹病度等級尺度或級距中位數取代NPE進行分析，有可能造成檢定力(power)下降的問題^(2, 3, 9, 29)。以 HB scale 為例，若以25 – 50%區間的中位數(37.5%)代表所有觀測值，則當實際觀測值為25%、30%或49%時，皆會被視為37.5%進行分析。此一概括性的處理，將導致病害資訊遺失，進而造成假設檢定時，虛無假設(例如： H_0 : Severity of treatment A = Severity of treatment B = ... =Severity of mock)為假但無法棄卻的問題(犯第二型錯誤)，此問題在樣本數

量有限時尤為嚴重⁽³⁾。在前述情況下，若評估者的判斷較不準確，在虛無假設為假的前提下，不論使用HB scale或NPE，成功拒絕虛無假設的機率相似；但若是較準確的評估者，一般而言使用NPE會具有較高的機率成功辨別處理間的差異⁽³⁾。

為了瞭解不同罹病度等級尺度結構對假設檢定結論的影響，蔣等人(2014)利用模擬資料，評估6種不同的罹病度等級尺度結構對於假設檢定時，犯第二型錯誤的機率的影響⁽⁹⁾。結果顯示，在罹病面積百分比小於5%時，若將罹病度等級尺度的級距設定為10%，在虛無假設為假的情況下，有較大的機率無法拒絕虛無假設；當罹病面積百分比介於20-50%時，使用HB scale的假設檢定結果都是最差的，亦即犯第二型錯誤的機率最高。若使用5%或10%為級距，且在罹病面積較小的範圍內增加細分的等級，該結構的罹病度等級尺度與NPE在進行假設檢定時，都具有良好的檢定力。因此，蔣氏尺度(‘Chiang’ scale)建議在0 – 10%的罹病面積範圍中，附加更細的尺度以提升解析度，而在罹病面積百分比在10%上時，則採10%級距設計罹病度等級尺度，以兼顧操作簡便性與統計檢定力(表二)⁽⁹⁾。

病害嚴重度指標(disease severity index; DSI)的估算與修正建議

在許多試驗中，研究者經常將前述的罹病度等級尺度資料，轉換為病害嚴重度指標(disease severity index, DSI)，以概括描述某樣本或處理結果的病害嚴重程度^(19, 20, 26, 28, 39)。因此，

表二、蔣氏尺度的病害分級方式與結構^(5, 9)

TABLE 2. The disease severity range and structure of the ‘Chiang Scale’^(5, 9)

Chiang Scale	Disease severity range	Midpoint
0	-	-
1	0 ⁺ - 0.1 %	0.05
2	0.1 ⁺ - 0.5 %	0.30
3	0.5 ⁺ - 1.0%	0.75
4	1.0 ⁺ - 2.0%	1.50
5	2.0 ⁺ - 5.0%	3.50
6	5.0 ⁺ - 10%	7.50
7	10.0 ⁺ - 20.0%	15.0
8	20.0 ⁺ - 30.0%	25.0
9	30.0 ⁺ - 40.0%	35.0
10	40.0 ⁺ - 50.0%	45.0
11	50.0 ⁺ - 60.0%	55.0
12	60.0 ⁺ - 70.0%	65.0
13	70.0 ⁺ - 80.0%	75.0
14	80.0 ⁺ - 90.0%	85.0
15	90.0 ⁺ - 100.0%	95.0

該指標除了提供作物在試驗條件下病害嚴重程度的量化依據，也被使用於比較不同處理間的效應⁽¹³⁾。

若探討量化順序尺度的罹病度資料，在過去的文獻中常見的DSI計算公式主要有兩種型式：

$$DSI = \sum(\text{罹病度的等級} \times \text{該級數中的樣本數量}) / \text{調查樣本總數} \quad \dots \dots \dots (1)$$

$$DSI\% = [\sum(\text{罹病度的等級} \times \text{該級數中的樣本數量}) / (\text{調查樣本總數} \times \text{最大罹病等級})] \times 100 \quad \dots \dots \dots (2)$$

兩式的主要差異在於後者將病害嚴重度指標，除以最大罹病等級標準化後，再乘以100轉換為百分比值。前者提供病害嚴重度指標的平均數值，後者則呈現標準化百分比。形式上，這些公式類似於教育測驗中的加權平均，例如 Grade Point Averages (GPA) 的計算方式。

然而，過去常見質化順序尺度的罹病度資料，例如：評估作物萎凋程度、病毒病徵的嚴重程度、或根瘤嚴重程度…等，被轉換成DSI後進行變異數分析(analysis of variance; ANOVA)的誤用。由於該類資料是依照病徵輕微至嚴重的情形，給予不同等級的描述，因此各等級的數字僅為排序。輕微至嚴重的等級，亦可以文字或符號表示排序，而該類資料的性質，較適合以無母數統計方法進行分析。相關方法過去已多有討論^(1, 29, 37)，在此不再贅述。

雖然DSI使用簡便，但其潛藏的系統性誤差需要留意。一般研究在分析此類資料時的做法包括：1. 使用公式(1)計算平均DSI後再進行ANOVA；2. 使用公式(2)計算DSI百分比值後，再進行ANOVA；3. 將罹病度數據，轉換為各罹病度等級所對應的級距中位數後，再進行ANOVA。前兩種方法，可能因為罹病度等級尺度的設計結構，或是公式既存的問題，而造成直接使用量化順序尺度，會導致使用「線性化」的方法處理對數尺度間距的資料，例如：認為等級2與等級3之間的差異與等級3和等級4之間的差異相近，進一步造成平均值的偏差⁽¹⁷⁾。在此，第三種方法是針對這類資料比較建議的分析方法^(1, 29)。

以Vieira等人的罹病度等級尺度為例，0 = 0%；1 = 0⁺-25%；2 = 25⁺-50%；3 = 50⁺-75%；4 = 75⁺-100%的病斑⁽³⁹⁾。舉例來說，假設在某處理中，調查樣本總數N = 4：其中1株屬於罹病度等級0、2株屬於等級1、1株屬於等級2，而等級3與4皆無樣本。依公式(1)計算，加權總和為(0 × 1 + 1 × 2 + 2 × 1 + 3 × 0 + 4 × 0)/4 = 4 / 4 = 1.0，亦即DSI = 1.0。接著按公式(2)， $DSI\% = [4 / (4 \times 4)] * 100 = 25\%$ 。

然而，在該處理中，DSI = 1.0表示該處理的平均罹病面積是介於0⁺-25%之間，因此，使用該罹病度的中位數12.5%進行後續分析較為恰當。若以公式(2)得出的25%進行後續分析，則會因為公式(2)忽略資料於該區間內的分布情形，可能導致系統性的高估罹病度⁽¹²⁾。

為了降低使用公式(2)轉換為以百分比表示的DSI時引入的

誤差，蔣等人(2017)提出，可將公式(1)計算的DSI，透過以下方式轉換為 $DSI_{Modified}$ ⁽¹²⁾：

$$DSI_{Modified} = Mid_Q + \frac{R}{N} \times (Mid_{Q+1} - Mid_Q) \quad \dots \dots \dots (3)$$

其中， $DSI_{Modified}$ 為修正後的病害嚴重度指標，該數值會與使用罹病度等級尺度的級距中位數所計算的平均值相同； Mid_Q 是商數對應的級距中位數；R為餘數，N為調查樣本總數，因此R/N為公式(1)計算出的DSI數值的小數部分； Mid_{Q+1} 為下一等級的級距中位數。再以Vieira等人(2010)的罹病度等級尺度為例⁽³⁹⁾，若公式(1)計算出的平均DSI = 1.6，使用公式(2)計算時， $DSI\% = \frac{1.6}{4} \times 100 = 40\%$ (其中，分母的4即為最大罹病等級)。然而，公式(1)計算的DSI介於1與2之間，若將該數值轉換為對應級距的中位數，DSI應該介於第一級的中位數(12.5%)與第二級的中位數(37.5%)之間較為合理。因此，以修正公式計算， $DSI_{Modified} = 12.5 + 0.6 \times (37.5 - 12.5) = 27.5\%$ ，而此數值，會與罹病度等級轉換成級距中位數後所計算的平均值相同⁽¹²⁾。此作法可以更細緻地反映不同等級之間的病害嚴重程度，避免罹病度高估的情形。

因此，傳統將罹病度換算為DSI雖然可以簡單的數字代表罹病度，但若未妥適考量病害等級尺度的結構，配合使用公式(2)計算DSI時，容易引入過多的誤差，導致高估罹病度。因此，在提高罹病度評估值的正確性方面，建議可使用標準病害面積圖(standard area diagrams)或電腦輔助訓練評估者，以增加評估者的評估能力。在評估病害時，建議使用NPE或結構適當的病害等級尺度(例如：'Chiang' scale)，並避免將罹病度的等級套用公式(2)後的數值進行後續分析。如需計算DSI%，可改以公式(3)進行計算，將能有效地解決高估罹病度的問題。

量化順序尺度資料的統計分析策略

前文已經簡單介紹量化順序尺度，在假設評估者都能夠準確地依照罹病度等級尺度進行評估的前提下，量化順序尺度的結構會影響到罹病度估計值的正確性。因此在評估罹病度時，一般而言使用NPE的方式為最佳。但若需使用罹病度等級尺度，以利快速的進行罹病度評估，蔣氏尺度(亦即罹病面積在10%以下時，將尺度細分成不同級距；罹病面積10%以上時，每10%為一個級距)的設計結構，能使得罹病度評估結果近似NPE的資料。在進行罹病度資料的統計分析時，可依照資料的特性，選擇以下合適的方法進行後續分析。

1. 中位數轉換後進行有母數分析 (parametric analysis)

Parametric analysis可翻譯作「參數分析」或「有母數分析」，本文採後者翻譯。依據美國植物病理學會期刊給予投稿人的建議以及過去的研究，NPE獲得的罹病度資料，在符合方法的假設前提下，使用ANOVA或線性混和模型(linear mixed model)分析較為恰當，雖然亦可使用無母數分析方法(non-parametric method)進行分析。而使用有母數分析方法時，可

能需要進行罹病度的轉換，因為變異數(variance)會隨平均值(mean)的改變而變化^(1, 5, 13)。

若是本文所提，屬於量化順序尺度的罹病度等級資料，在分析前，建議將各等級的資料轉換為該等級對應的級距中位數後，再進行有母數分析(parametric analysis)。不建議以前述公式(2)將罹病度資料轉換為DSI(%)後進行分析，因為該方法會導致高估罹病度的問題。如需轉換，可以利用公式(3)將公式(1)的DSI轉換為DSI_{Modified}。轉換後的DSI_{Modified}，會與直接使用中位數轉換的數值相同，之後再進行有母數分析。使用有母數分析方法的優點是統計檢定力(power)較佳，但缺點在於資料特性需要符合方法的諸多假設前提，因此在進行分析前，需要先確認資料的特性。如不將罹病等級轉換為中位數，而是直接以罹病等級資料進行分析，可參考下文所提的幾個方法。

2.無母數統計分析方法(Non-parametric methods)

針對資料無法滿足常態分布假設或資料為不連續變數的情況，例如：利用量化順序尺度與質化順序尺度所收集的罹病度資料，無母數分析方法例如：Mann - Whitney U test (對應的有母數分析方法為t-test)、Kruskal - Wallis test(對應的有母數分析方法為One-way ANOVA)等，可以用以分析此類型的資料。此類方法並未預先假設資料的分布形態，而是將資料排序後，再進行中位數的比較。無母數分析方法的優點包括：1.無須滿足資料符合常態分佈與變異數均質性的假設；2.適用於順序尺度資料；3.適用於樣本數較小的研究設計。然而，無母數分析方法的缺點，在於檢定力與有母數分析方法相比較弱。因此在可正確精準量化病徵面積的情況下，仍建議使用有母數方法進行罹病度資料的統計分析^(1, 21, 29)。

3.比例勝算模型(Proportional Odds Model)

比例勝算模型(proportional odds model, POM)屬於有序邏輯斯回歸模型(ordered logistic model)的一種，適合用於分析順序尺度資料。過去也曾有許多研究，將該模型應用於比較不同處理對罹病度的影響，或是用於建立預測模型^(14, 18, 23, 27)。與其他有母數分析方法不同處，在於此模型不需要將罹病等級資料轉換為對應級距的中位數，而是直接分析各罹病度等級累積機率的變化。模型如下：

$$\text{logit}[P(Y \leq j)] = \log \left(\frac{P(Y \leq j)}{1 - P(Y \leq j)} \right) = \alpha_j + \beta x$$

其中， j 為罹病度的等級； $P(Y \leq j)$ 為罹病度小於等級 j 的機率； α_j 為截距，參數 β 用以解釋變數 x 對第 j 級罹病度之累積對數機率值(cumulative log odds)的影響。蔣等人(2020)的模擬研究結果顯示，使用NPE再配合有母數分析方法，一般來說具有最佳的統計檢定力⁽¹⁴⁾。而當罹病度低於 40%時，使用POM直接分析罹病度等級，或是將罹病度等級轉換為該級對應的級距中位數後再進行有母數分析，兩者都有不錯的檢定力。而在罹病度

小於10%，且評估值準確度較低時，直接將罹病度等級資料以POM 分析的檢定力，優於中位數轉換後再進行有母數分析的檢定力。因此，當罹病面積低於40%時，例如病害發生初期，或是比較不同處理對於罹病度的影響時，建議可以POM進行分析；而罹病面積高於40%時，POM的檢定力則與其他方法沒有太大差異⁽¹⁴⁾。

除了前述的優點以外，使用POM進行處理間的比較時，應留意樣本數量是否足夠進行分析，否則將無法得到較佳的結果⁽³⁷⁾。而模擬研究的結果顯示，在多數的情形下，每個試驗單位至少需要15個以上的樣本，較能適用該方法⁽¹⁴⁾。

4.存活分析(Survival Analysis)

前述的各種方法中，我們常使用量化順序尺度資料，以罹病度等級取代NPE，記錄病害的嚴重程度。在進行統計分析時，則將罹病度等級，轉換為該等級的罹病面積級距中位數後，再進行有母數分析。然而，將罹病度等級轉換為級距中位數的方法，主要是為了避免在分析時，高估罹病度或導入過多誤差，而將中位數視為該區間的估計值。如果在罹病面積級距太大的情形下，將罹病度等級轉換為中位數，亦無法完全代表落於該區間的罹病面積資料。在此前提之下，也可能引入誤差而影響後續的分析結果。因此，過去曾有學者提出使用存活分析(survival analysis)的方法，進行植物保護與雜草相關研究的數據分析^(8, 22, 32, 33, 34, 36)。

存活分析的資料種類，包括左設限資料(left-censored data)、右設限資料(right-censored data)，以及區間設限資料(interval-censored data)。舉例來說，在所有的供試目標，經歷研究想要探討的事件或現象(例如：病害發生、植株死亡…等)之前，研究就已經結束，或者供試目標因其他原因被排除在研究之外，此時就稱為「設限(censoring)」，而前述的案例中收集的資料，則為右設限資料。若在供試目標納入研究前，就已經發生欲探討的事件或現象，但發生的時間不確定，此種情形稱為左設限。而當我們知道研究欲探討的事件發生在某個時間段內，但確切的時間不確定時，此種情形稱為區間設限。雖然存活分析的主要目的是探討某事件發生所需的時間(time to event)，但根據這個概念，也可以將罹病度等級資料(例如 HB scale)視為區間設限資料，亦即我們對其數值有部分了解，但並非確切量測值。由此可套用存活分析的方法，來比較不同處理對於罹病度的影響^(6, 15, 38)。

蔣等人(2024)研究指出，存活分析的方法在樣本罹病度不是呈現常態分佈時，幾乎都能比中位數轉換法的分析，具有較佳的統計檢定力。在某些情況下，中位數轉換法的樣本數，需要存活分析樣本數的4-5倍，才能達到相同的檢定力。而當平均罹病度增加時，不論使用何種病害評估方法(NPE、HB scale或蔣氏尺度)，相較於存活分析，中位數轉換法需要2倍的樣本數，才能達到與存活分析近似的檢定力。因此，若使用量化順

序尺度方法進行病害嚴重程度的評估時，建議使用存活分析進行後續的統計分析與處理比較，將具有最佳的穩定性與檢定力⁽¹⁵⁾。

結論與建議

本文針對植物病害嚴重度的量化評估進行一系列的探討，從病徵面積的評估問題出發，回顧並探討如何優化罹病度等級尺度的設計結構，以及罹病度資料後續的統計分析方法。透過過去的研究成果，針對如何提升罹病度的評估精準度(precision)與正確度(accuracy)(合稱準確度)，以及建議的罹病度資料統計分析方法，歸納如下。

1. 提升罹病度評估準確度的作法

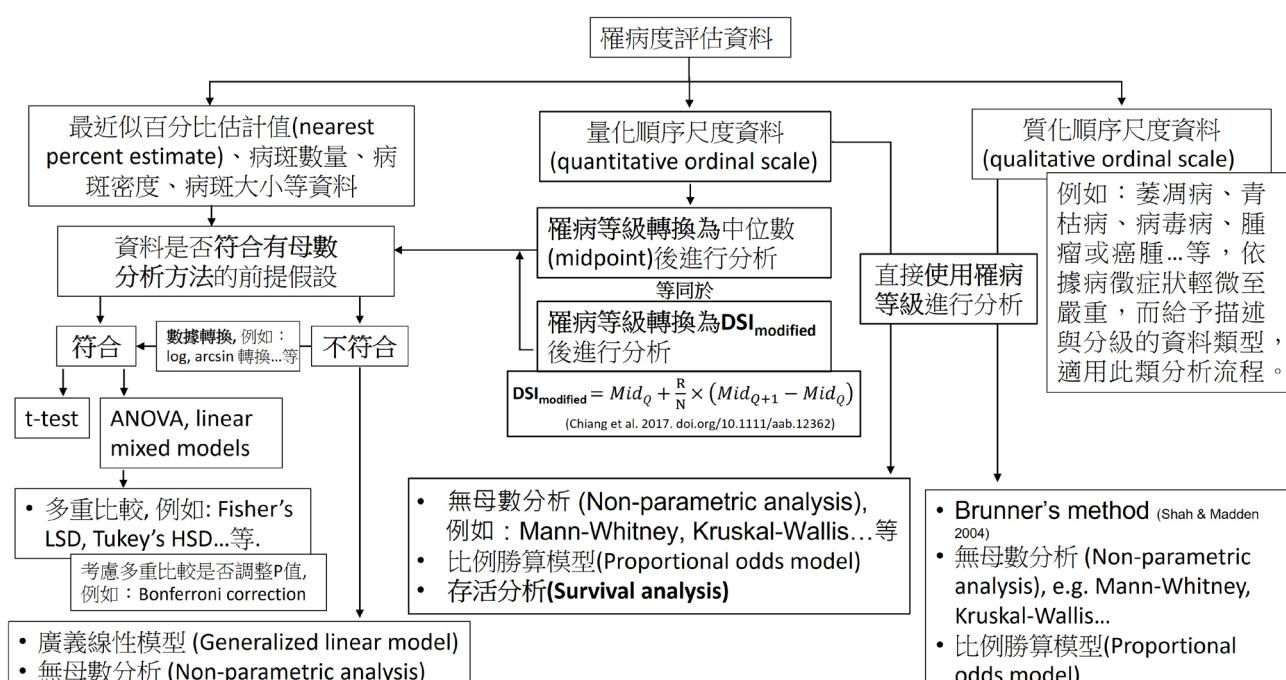
根據過去的研究成果，Bock等人提出幾點建議，以提昇病害評估時的準確度⁽⁵⁾。首先，依據不同的病害系統、研究的需求、以及可獲得的資源，來選擇最適合的罹病度等級尺度。使用NPE的評估方式雖然可提供最大量的病害資訊，但若無法評估NPE，評估者也需要先瞭解量化順序尺度的特性與應用方式，並選擇結構合適的罹病度等級尺度進行病害評估。例如，在田間病害調查時可使用蔣氏尺度(表二)，亦即罹病面積在10%以下，多細分幾個等級；罹病面積在10%以上，則以10%為一個級距。

其次，評估者對於所評估的病害種類、罹病度等級尺度、以及評估時的誤差來源，需要有一定程度的瞭解。因此，充分的指導與訓練評估者，將有助於提升病害評估的準確度。在訓練評估者時，可以利用標準病害面積圖、電腦程式輔助、或是實際樣本的方式，多能有效提升準確度。評估者在接受訓練後，可進一步利用已知的資料進行測試，如此可以瞭解評估者在進行病害評估時的準確度和一致性，以及能否在接受訓練後，提升病害評估的準確度。評估者也需要累積病害評估相關的經驗，才能穩定且準確地進行病害評估。

最後，若在可能的情況下，使用最少量的評估者進行評估，或是在評估不同處理的試驗時，將評估者隨機分配到不同試驗單位(experimental units)中，如此能有助後續的統計分析中，辨識並控制評估者間的誤差對研究造成的影響。若樣本取得困難，亦可透過增加評估重複次數提高資料準確度，而兩次觀測即足以大幅降低評估誤差^(10, 11)。

2. 罹病度資料的統計分析建議

本文著重於探討量化順序尺度的資料取得與分析，在實務操作上，若無法直接評估NPE，可以依照前述建議進行病害評估後，再依據罹病度資料的特性與適合的方法進行統計分析(圖二)。因此在分析前，可先將罹病度等級轉換為該等級對應的罹病面積級距中位數後再進行有母數分析；若要轉換為病害嚴重度指標(DSI)，可利用前述的公式(1)或公式(3)進行計算，



圖二、罹病度資料的建議分析流程。

Fig. 2. The workflow for analyzing disease severity data.

因為該 $DSI_{Modified}$ 值會等同於中位數轉換後的數值，之後再進行有母數分析。若資料在轉換後，不符合有母數分析的假設前提，可以進行數據轉換，測試是否符合有母數分析的假設前提後，再進行有母數分析與假設檢定。

若直接使用罹病度等級進行分析，可以考慮使用無母數分析、比例勝算模型或存活分析方法進行假設檢定。需要留意的是，在使用無母數分析時，其檢定力可能較有母數分析弱；而使用比例勝算模型時，則應該有足夠的樣本數(理想上>30樣本)，以利進行處理間的比較與假設檢定。依照目前的研究成果而言，若在研究中使用合適的罹病度等級尺度，存活分析應該是分析量化順序尺度取得的罹病度資料的最佳方法。因為該方法不論在檢測低罹病度或是存在較大的評估者誤差的資料上，都能夠有一定優於NPE資料的檢定力，因此未來可再深入探討該方法應用於植物病理相關研究資料的分析上。

綜合前述各種分析方法之適用條件與研究建議，為協助非統計背景之植物病理相關研究人員能依據資料特性選擇合適的統計分析方法，本文特將不同樣態的病害量化資料所對應的建議分析策略整理如下(表三、圖二)。表三彙整樣本數多寡、資料型態(如NPE、順序等級尺度)、是否符合常態分布、是否為設限資料等常見情形，對應可使用之統計方法，並提供簡要說明，期能作為實務應用之參考依據。

後 記

本文為整理2025中華民國植物病理學會年會中，「植物病害量化評估之理論與實務」工作坊介紹之內容。

引用文獻

1. APS. 2025. Information for authors. *Phytopathology*. American Phytopathological Society [WWW document]. URL <https://apsjournals.apsnet.org/page/authorinformation#statistics> [accessed on 22 May 2025].
2. Bock, C. H., Gottwald, T. R., Parker, P. E. and Cook, A. Z., Ferrandino, F., Parnell, S., and van den Bosch, F. 2009. The Horsfall-Barratt scale and severity estimates of citrus canker. *Eur. J. Plant Pathol.* 125: 23 – 38.
3. Bock, C. H., Gottwald, T. R., Parker, P. E., Ferrandino, F., Welham, S., van den Bosch, F., and Parnell, S. 2010. Some consequences of using the Horsfall-Barratt scale for hypothesis testing. *Phytopathology* 100:1031-1041.
4. Bock, C. H., Poole, G. H., Parker, P. E., and Gottwald, T. R. 2010. Plant disease severity estimated visually, by digital photography and image analysis, and by hyperspectral imaging. *Crit. Rev. Plant Sci.* 29: 59 – 107.
5. Bock, C. H., Chiang, K. S. and Del Ponte, E. M. 2022. Plant disease severity estimated visually: a century of research, best practices, and opportunities for improving methods and practices to maximize accuracy. *Trop. Plant Path.* 47:25 – 42.
6. Bogaerts, K., Komárek, A., and Lesaffre, E. 2017. *Survival Analysis with Interval-Censored Data: A Practical Approach with Examples in R, SAS, and BUGS*. Chapman & Hall/CRC, Boca Raton, 616 pp.
7. Campbell, C. L., and Madden, L. V. 1990. *Introduction to Plant*

表三、罹病度資料的統計分析建議方法對照表

TABLE 3. Recommended statistical analysis methods for disease severity data

資料特性或條件	統計分析建議方法	備註說明
最近似百分比估計值數據(nearest percent estimate; NPE)、病斑數量、病斑密度、病斑大小等資料	使用有母數分析方法。成對比較可使用Student's <i>t</i> -test；三個處理以上的比較可使用ANOVA、線性混合模型等。	建議先檢驗資料的常態性與變異數齊一性，可視需要進行適當的資料轉換。
量化與質化順序尺度資料且樣本數充足(每處理 ≥ 15 樣本)	比例勝算模型(Proportional Odds Model)。	無須轉換為中位數，可直接分析等級的機率變化。
使用 HB scale、Chiang scale 等量化順序尺度資料	存活分析(Interval-censored Survival Analysis)。	特別適用於低罹病度或存在評估誤差時，檢定力優於其他方法。
量化順序尺度資料轉換為對應的罹病面積中位數資料	罹病度資料轉換為等級對應的罹病面積中位數或使用公式(3)： $DSI_{Modified}$ 轉換後，再評估能否依照有母數分析方法(例如：ANOVA、線性混合模型等)進行分析。	不建議使用公式(2)之 $DSI(\%)$ ，以避免系統性高估罹病度。
資料不符合常態分布或樣本數較小、質化順序尺度資料	使用無母數分析方法。成對比較可使用Mann-Whitney U 檢定；三組以上的處理比較可使用Kruskal-Wallis 檢定。	假設前提需求較少，適用於順序資料，但檢定力較弱。

- Disease Epidemiology. John Wiley & Sons, New York, 532 pp.
8. Chen, M., Brun, F., Raynal, M., and Makowski, D. 2019. Timing of grape downy mildew onset in Bordeaux vineyards. *Phytopathology* 109:787-795.
 9. Chiang, K. S., Liu, S. C., Bock, C. H., and Gottwald, T. R. 2014. What interval characteristics make a good categorical disease assessment scale? *Phytopathology* 104:575-585.
 10. Chiang, K. S., Bock, C. H., El Jarroudi, M., Delfosse, P., Lee, I. H., and Liu, H. I. 2016. Effects of rater bias and assessment method on disease severity estimation with regard to hypothesis testing. *Plant Pathol.* 65:523-535.
 11. Chiang, K. S., Bock, C. H., Lee, I. H., El Jarroudi, M., and Delfosse, P. 2016b. Plant disease severity assessment - how rater bias, assessment method and experimental design affect hypothesis testing and resource use efficiency. *Phytopathology* 106:1451-1464.
 12. Chiang, K.S., Liu, H. I., and Bock, C.H. 2017. A discussion on disease severity index values. Part I: warning on inherent errors and suggestions to maximize accuracy. *Ann. Appl. Biol.* 171:139-154.
 13. Chiang, K. S., Liu, H. I., Tsai, J. W., Tsai, J. R., and Bock, C. H. 2017. A discussion on disease severity index values. Part II: using the disease severity index for null hypothesis testing. *Ann. Appl. Biol.* 171:490-505.
 14. Chiang, K. S., Liu, H. I., Chen, Y. L., El Jarroudi, M., and Bock, C. H. 2020. Quantitative ordinal scale estimates of plant disease severity: comparing treatments using a proportional odds model. *Phytopathology* 110:734-743.
 15. Chiang, K. S., Chang, Y. M., Liu, H. I., Lee, J. Y., El Jarroudi, M., and Bock, C. H. 2024. Survival analysis as a basis to test hypotheses when using quantitative ordinal scale disease severity data. *Phytopathology* 114:378-392.
 16. Cobb, N. A. 1892. Contribution to an economic knowledge of the Australian rusts (*Uredinae*). *Agric. Gazt. (NSW)* 3:60.
 17. Forbes, G. A., and Korva, J. T. 1994. The effect of using a Horsfall-Barratt scale on precision and accuracy of visual estimation of potato late blight severity in the field. *Plant Pathol.* 43:675-682.
 18. Fu, L. Y., Wang, Y. G., and Liu, C. J. 2012. Rank regression for analyzing ordinal qualitative data for treatment comparison. *Phytopathology* 102:1064-1070.
 19. Gafni, A., Calderon, C. E., Harris, R., Buxdorf, K., Dafa-Berger, A., Zeilinger-Reichert, E., and Levy, M. 2015. Biological control of the cucurbit powdery mildew pathogen *Podosphaera xanthii* by means of the epiphytic fungus *Pseudozyma aphidis* and parasitism as a mode of action. *Front. Plant Sci.* 6:132.
 20. Ghose, L., Neela, F. A., Chakravorty, T. C., Ali, M. R., and Alam, M. S. 2010. Incidence of leaf blight disease of mulberry plant and assessment of changes in amino acids and photosynthetic pigments of infected leaf. *Plant Pathol. J.* 9:140-143.
 21. Gibbons, J. D., and Chakraborti, S. 2011. Nonparametric Statistical Inference (5th ed.). Chapman & Hall/CRC, New York, 650 pp.
 22. Gottwald, T. R., Hall, D. G., Kriss, A. B., Salinas, E. J., Parker, P. E., Beattie, G. A. C., and Nguyen, M. C. 2014. Orchard and nursery dynamics of the effect of interplanting citrus with guava for huanglongbing, vector, and disease management. *Crop Prot.* 64:93-103.
 23. Henderson, D., Williams, C. J., and Miller, J. S. 2007. Forecasting late blight in potato crops of southern Idaho using logistic regression analysis. *Plant Dis.* 91: 951-956.
 24. Horsfall J. G. and Barrat, R. W. 1945. An improved grading system for measuring plant disease. *Phytopathology* 35: 655 (Abstract).
 25. Hunter, R. E., and Roberts, D. D. 1978. A disease grading system for pecan scab. *Pecan Qtly.* 12:3-6.
 26. Koitabashi M. 2005. New biocontrol method for parsley powdery mildew by the antifungal volatiles-producing fungus Kyu-W63. *J. Gen. Plant Pathol.* 71: 280-284.
 27. Landschoot, S., Waegeman, W., Audenaert, K., Haesaert, G., and Baets, B. 2013. Ordinal regression models for predicting deoxynivalenol in winter wheat. *Plant Pathol.* 62:1319-1329.
 28. Lazarovits, G., Hill, J., Patterson, G., Conn, K. L., Crump, N. S. 2007. Edaphic soil levels of mineral nutrients, pH, organic matter, and cationic exchange capacity in the geocalosphere associated with potato common scab. *Phytopathology* 97:1071-1082.
 29. Madden, L. V., Hughes, G., and van den Bosch, F. 2007. The Study of Plant Disease Epidemics. APS Press, St. Paul, 421 pp.
 30. Nita, M., Ellis, M. A., and Madden, L. V. 2003. Reliability and accuracy of visual estimation of *Phomopsis* leaf blight of strawberry. *Phytopathology* 93: 995 – 1005.
 31. Nutter, F. W. and Esker, P. D. 2006. The role of psychophysics in phytopathology: The Weber-Fechner law revisited. *Eur. J. Plant Pathol.* 114:199-213.
 32. Ojiambo, P. S., and Kang, E. L. 2013. Modeling spatial frailties in survival analysis of cucurbit downy mildew epidemics. *Phytopathology* 103: 216-227.
 33. Ojiambo, P. S., and Scherm, H. 2005. Survival analysis of time

- to abscission of blueberry leaves affected by *Septoria* leaf spot. *Phytopathology* 95: 108-113.
34. Onofri, A., Piepho, H. P., and Kozak, M. 2019. Analyzing censored data in agricultural research: A review with examples and software tips. *Ann. Appl. Biol.* 174:3-13.
35. Redman C. E., King, E. P., and Brown I. F., Jr. 1968. Tables for Converting Barratt and Horsfall Rating Scores to Estimated Mean Percentages. Elanco Products, Indianapolis, 8 pp.
36. Scherm, H., and Ojiambo, P. S. 2004. Applications of survival analysis in botanical epidemiology. *Phytopathology* 94:1022-1026.
37. Shah, D. A., and Madden, L. V. 2004. Nonparametric analysis of ordinal data in designed factorial experiments. *Phytopathology* 94:33-43.
38. Turnbull, B. W. 1976. The empirical distribution function with arbitrarily grouped, censored and truncated data. *J. R. Stat. Soc. Series B.* 38: 290 – 295.
39. Vieira, R. F., Paula Jr., T. J., Teixeira, H., and Carneiro, J. E.D.S. 2010. White mold management in common bean by increasing within-row distance between plants. *Plant Dis.* 94:361-367.
- and accuracy of disease severity assessments, and introduces recent advances in statistical methods for analyzing quantitative ordinal scale data, aiming to enhance researchers' understanding of the collection and analysis of such data in plant protection research.

Keywords: plant disease quantification, disease severity, quantitative ordinal scale, proportional odds model, survival analysis

ABSTRACT

Kuo-Szu Chiang and Cheng-Fang Hong*. 2025. Optimizing ordinal scales for plant disease severity assessment and advances in their statistical analysis. *J. Plant Med.* 67(2): 69-78.

*Corresponding author, E-mail: cfhong@nchu.edu.tw

Plant disease quantification is crucial for crop disease management, breeding for disease resistance, and the development of disease forecasting systems, as well as for related research and applications. Common approaches to disease quantification include evaluating disease prevalence, incidence, and severity. Traditionally, disease severity has been defined as the percentage of symptomatic area on plant tissues or organs, with the nearest percent estimate (NPE) often used as a measure. However, not all raters can quickly, precisely, and accurately assess NPE. Thus, many studies have proposed the use of quantitative ordinal scales for estimating disease severity. One example is the Horsfall-Barratt scale, which has been widely applied in studies of various plant diseases. Nonetheless, due to its nonlinear scale design, this type of scale has been criticized for introducing bias that may reduce statistical power in subsequent analyses. This article reviews past studies, discusses how to optimize the structure of quantitative ordinal scales to improve the precision

