Reducing over-diagnosis and over-treatment by improving their criteria and stratifying them

Huw Llewelyn Aberystwyth University, Wales, UK. hul2@aber.ac.uk

Objectives

Many patients with the historical diagnosis of 'dropsy' (swelling of the lower body due to fluid) were later 'stratified' into those with the rapid irregular pulse (i.e. atrial fibrillation treatable by digitalis), high levels of protein in the urine and so on. Forty years ago it was postulated that small amounts of albumin in the urine of diabetic patients might predict later nephrotic syndrome (one cause of 'dropsy'). This led to routine urine testing for albumin in diabetics and if it was raised, treatment with an angiotensin-converting-enzyme inhibitor (ACEI) or an angiotensin receptor blocker (ARB). If the patient is known to have diabetes mellitus then it is not certain that the cause is early kidney damage from diabetes that can be prevented. To make it more probable that this is the cause (and not to over-diagnose), it is customary to show that other causes (e.g. poorly controlled hypertension, nephritis, heart failure, recent exercise, prolonged standing, UTI or a borderline result) are improbable or excluded or treated. If this is not done, it may be assumed falsely that early diabetic nephropathy is probable, resulting in over-diagnosis and over-treatment. The object of this presentation is to show how this type of reasoning can be used to reduce over-diagnosis based on all numerical tests.

Method

Data were analyzed from the IRMA2 trial that studied 600 patients with type 2 diabetes mellitus, treated hypertension and albuminuria [1]. Those with self-limiting conditions (e.g. prolonged standing) and diagnoses needing other treatments (e.g. a nephritis) were excluded by using 'reasoning by probable elimination' [2]. This process is illustrated in Table 1. Patients were then randomized to placebo (200 patients) or the angiotensin receptor blocker irbesartan (200 patients on 150mg od and 200 on 300mg od). The patients were 'stratified' into groups with initial albumin excretion ranges (AER) of 20 to 40mcg/min, 41 to 80, 81 to 120 and 121 to 200mcg/min. The proportion in each group developing 'nephropathy' (an AER > 200mcg/min) after two years were then plotted in a histogram as shown in Figure 1.

Table 1:
Reasoning by elimination when AER > 20mcg/min in a set of patients with Type 2 diabetes mellitus

? Diabetic nephropathy in future
(frequent because BP normal, no blood in urine, no dipstick signs of UTI, normal heart size, morning urine collected after being supine overnight, AER > 40mcg/min, complete urine collection and HbA1c not high)

Causes of over-diagnosis:
? No hypertension (infrequent because of normal BP)*
? No congestive cardiac failure (infrequent as normal heart size)
? Urinary tract infection (infrequent as no dipstick signs of UTI)
? No exercise induced proteinuria (infrequent as early morning urine used after lying supine overnight)
? No orthostatic proteinuria (infrequent as early morning urine used (see above*)
? No spurious albuminuria (infrequent because AER > 40mcg/min)
Causes of under-diagnosis:
? Not dehydrated or Incomplete collection (history, examination, creatinine etc)
Causes of rapid worsening or poor treatment response due to other active processes e.g:
? No nephritis (infrequent as no blood in urine)
? No poor diabetic control (infrequent as HbA1c not high)

Abstract

Objectives: The aim of this presentation is to demonstrate the role evidence-based diagnostic criteria and their stratification in stopping and preventing future over-diagnosis and over-treatment. Ideal diagnostic criteria do not exclude patients who benefit and do not label people who have no prospect of benefiting from the actions directed at that diagnosis. It is often necessary to 'stratify' those with a diagnosis further into subsets within which the frequency of those benefiting from each treatment is high to avoid over-treatment..
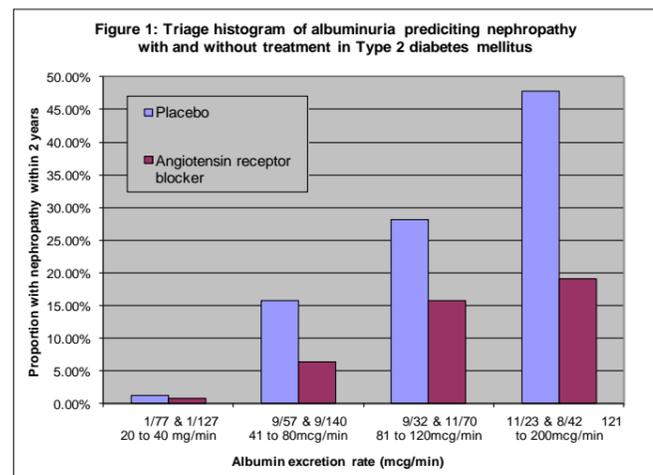
Methods: The principles will be illustrated with the data from a randomized controlled trial on the effect of an angiotensin receptor blocker to treat patients with type 2 diabetes mellitus who have diabetic albumuinuria (D-A) to prevent them developing diabetic nephropathy and then nephrotic syndrome (one of the diagnoses formulated historically by stratifying 'dropsy'). The severity of the D-A was stratified into ranges and also stratified by using other findings such as blood pressure.

Results: When patients with D-A were stratified into different ranges, the absolute proportions benefiting (and the number needed to treat) were seen to vary widely between each stratified group. The patients with D-A whom were stratified into different sub-diagnoses based on other features (such as very high blood pressure) were excluded from the trial to avoid misdiagnosis.

Conclusion: It is essential to understand the principles of evidence based diagnostic criteria in order to stop over-diagnosis and over-treatment and to prevent it in future.

Results

The numbers removed from the trial because of the exclusion criteria described in table 1 were not recorded. In the remaining patients, there was a marked difference in the proportion developing 'nephropathy' in the four stratified groups in figure 1. In patients with an AER between 20 and 40mcg/min only 1/77 developed 'nephropathy' on placebo and 1/127 did so on the two different doses of irbesartan. In other words, when the AER is less than 40mcg/min, the number needed to treat in order to prevent one patient getting 'nephropathy' is 1/(1/77 – 1/127) = 195. It is only those 75% of patients with an AER over 40mcg/min that have a substantial risk of developing nephropathy and where treatment can reduce the risk in a helpful way.

Figure 1: Triage histogram of albuminuria predicting nephropathy with and without treatment in Type 2 diabetes mellitus

Discussion

The end-point chosen in the IRMA 2 trial was biochemical. No symptoms or measures of well-being were included. However, the principle can be applied to any type of patient's outcome, the purpose of this presentation being to outline the methods required to reduce over-diagnosis. The ideal histogram columns for a highly predictive AER test would show a steep rise from 0% of patients with a poor outcome at a low AER values to 100% with a poor outcome at high levels. This would mean that a low AER predicts a 'good outcome' and a high AER a 'poor outcome'. In figure 1, the control (blue) column

does predict a good outcome at a low level (i.e. an AER < 40mcg/min). However, if we assume that patients with an AER of 20 to 40mcg/min have 'Diabetic albuminuria' and have to be treated, then we will over-diagnose and over-treat this group. If we raise the cut-off point from 20 to 40mcg/min, then we will reduce the number diagnosed by 25% and no patients will suffer from being excluded, thus reducing over-diagnosis. Many patients with a raised AER due to other reasons (e.g. UTI, prolonged standing, etc) will have been excluded already as shown in table 1. However, the technique shown here of also stratifying patients according to the severity of a predictive test result reduces over-diagnosis further and should be applied to all diagnoses and treatment selection criteria based on numerical test results.

New research to reduce over-diagnosis by stratification

Figure 1 shows that even when the AER is between 120 and 200mcg/min, only 48% of patients in the control group develop 'nephropathy' (an AER over 200mcg/min) after 2 years. If the initial AER had been a powerful predictor of the AER after 2 years, then at an AER of 200mcg/min, the control column would indicate that 100% of patients had 'nephropathy', so that the columns rose more steeply . This suggests that some other factor is operating to reduce the AER or preventing its further rise. One possibility is that the increased attention paid to patients during the trial is leading to subtle improvement in diabetic control over 2 years. However, if 'nephropathy' had been frequent following low levels of AER (making the curve shallower) this might suggest loss of diabetic control or some other undiagnosed negative effect in some patients. These hypotheses would have to be explored in a separate study in an attempt to improve diagnostic accuracy by stratifying patients according to other factors which are not recognized at present

Improving treatment response

The benefit from treatment is represented by the difference in the proportions with a poor outcome in the control (blue) and treatment (red) columns in figure 1. For example, at an AER between 20 and 40mcg/min, only 1.25% – 0.75% = 0.5% benefit but at an AER between 120 and 200mcg/min 48%-19% = 29% benefit. The proportion benefiting can also be improved by not treating those in whom something else is more effective e.g. those with poor diabetic control or some other yet undiscovered process. It is also important to ensure that the treatment being assessed is given optimallye.g. at the right dose etc of course.

The effect of shared decision making

A diagnosis is the title to a number of predictions about the past, present and future with or without the effect of interventions. As shown in figure 1, the proportion of patients within a group benefiting varies. However, each group will suffer the same proportion of harmful effects. The decision to accept or reject a treatment depends on how a patient views these benefits and harms and the probability of them happening. It is therefore essential for the probability of benefit for each value of AER to be made clear as shown in figure 1.

Conclusion

Over-diagnosis and over-treatment can only be prevented by a clear understanding of the logic of diagnostic criteria and how the findings that form these criteria can be stratified to improve prediction.

References

1. Llewelyn D E H, Garcia-Puig, J. How different urinary albumin excretion rates can predict progression to nephropathy and the effect of treatment in hypertensive diabetics. JRAAS 2004, 5; 141-5.
2. Llewelyn H, Ang AH, Lewis K, Abdullah A. The Oxford Handbook of Clinical Diagnosis, 3° edition. Oxford University Press, Oxford, 2014.