University of Sheffield | Sheffield Centre For Health & Related Research

HEALTH ECONOMICS & DECISION SCIENCE

Discussion Paper Series

HEDS Discussion Paper 24.04

Title: The cost-effectiveness of using waking salivary cortisol in the diagnosis of adrenal sufficiency

Authors: Simon Dixon, Jen Lewis, Miguel Debono, Charlotte J Elder, Sharon Caunt, Richard M Jacques, John Newell-Price, Martin J Whitaker, Brian Keevil, Richard J Ross

Corresponding author: Simon Dixon, Professor of Health Economics, SCHARR, University of Sheffield, Regent Court, 30 Regent Street, Sheffield, S1 4DA, UK. Email: @professorsimondixon@yahoo.com

Disclaimer:
This series is intended to promote discussion and to provide information about work in progress.
The views expressed in this series are those of the authors.
Comments are welcome, and should be sent to the corresponding author.

The cost-effectiveness of using waking salivary cortisol in the diagnosis of adrenal sufficiency

Simon Dixon, Jen Lewis, Miguel Debono, Charlotte J Elder, Sharon Caunt, Richard M Jacques, John Newell-Price, Martin J Whitaker, Brian Keevil, Richard J Ross

Affiliations

Simon Dixon

School of Medicine and Population Health, University of Sheffield, Sheffield, United Kingdom

Jen Lewis

School of Medicine and Population Health, University of Sheffield, Sheffield, United Kingdom

Miguel Debono

Department of Endocrinology, Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, United Kingdom

School of Medicine and Population Health, University of Sheffield, Sheffield, United Kingdom

Charlotte J Elder

School of Medicine and Population Health, University of Sheffield, Sheffield, United Kingdom

Sharon Caunt

School of Medicine and Population Health, University of Sheffield, Sheffield, United Kingdom

Richard M Jacques

School of Medicine and Population Health, University of Sheffield, Sheffield, United Kingdom

John Newell-Price

School of Medicine and Population Health, University of Sheffield, Sheffield, United Kingdom

Martin J Whitaker

School of Medicine and Population Health, University of Sheffield, Sheffield, United Kingdom

Brian Keevil

Department of Biochemistry, Manchester University NHS Foundation Trust, Manchester, United Kingdom

Manchester Academic Health Sciences Centre, Manchester, United Kingdom

Richard J Ross

School of Medicine and Population Health, University of Sheffield, Sheffield, United Kingdom

Abstract

Adrenal Insufficiency (AI) is a deficiency in the stress hormone cortisol, and when untreated it can lead to life-threatening complications. Clinically, AI can present with non-specific signs and symptoms such as fatigue, muscle aches and abdominal pain, often resulting in delayed diagnosis. Currently, the Short Synacthen test (SST) is the reference standard for diagnosing AI, but requires venepuncture and a 2-3 hour hospital attendance. A new home-based test is available that uses a patient's waking salivary cortisone (WSC) levels as a surrogate marker of that collected by the SST. As well as being more convenient for patients, the test only costs £19 per patient compared to the £184 for the SST.

A recent diagnostic accuracy study undertaken in Sheffield compared SST and WSC and produced negative and positive prediction values for AI of 96% (95% CI; 90 to 99) and 95% (95% CI; 87 to 99), respectively. However, despite the potential costs saving of £165 per patient if WSC was to directly replace SST, this is not considered realistic given the slightly inferior diagnostic performance of WSC relative to the SST. Consequently, it is more likely to be adopted as part of a two-stage diagnostic strategy, with WSC used to screen for AI and the hospital-based SST used to test for AI in patients with an equivocal WSC result.

In this study, we report an economic evaluation of a two-stage diagnostic strategy from the NHS perspective and with a time horizon bounded by the generation of a non-equivocal biochemical diagnosis. This is undertaken by formulating a decision tree to look at costs and diagnostic performance for the two approaches, which is populated from the Sheffield diagnostic accuracy study and local unit costs.

The economic results for the primary analysis show that a two-stage diagnostic strategy costs £103 per patient less than the current diagnostic strategy (or 58% less). This is associated with a worse diagnostic performance with 7 out of 220 patients (3.2%) receiving a false diagnosis (four false positive and three false negative diagnoses). When the secondary analyses are considered, the use of alternative diagnostic cutoffs for SST has little impact on the results. The use of a societal perspective increases cost savings to £127 per patient (or a 59% reduction relative to current care).

The two-stage strategy is considered to be a valuable alternative to the current diagnostic process, due to its cost savings, that are associated with improved patient experience and quicker diagnosis for the vast majority of patients. Other benefits include the enhanced resilience of the WSC testing process due to the avoidance of hospital visits (which becomes even more important during health service crises, including pandemics) and the need to procure synacthen. Further studies should look to build on our findings, by adopting the proposed two-stage diagnostic pathway in the routine clinical practice of other hospitals.

Introduction

Adrenal Insufficiency (AI) is a deficiency in the essential stress hormone cortisol, secreted by the adrenal glands. When untreated, adrenal insufficiency can result in an adrenal crisis, with life-threatening cardiovascular collapse and a mortality rate of around 6% (Rushworth et al, 2019). Annually 6-8% of patients living with AI experience a crisis (Rushworth et al, 2019). Clinically, AI can present with non-specific signs and symptoms such as fatigue, muscle aches and abdominal pain, often resulting in delayed diagnosis, risking death through an adrenal crisis.

Currently, the Short Synacthen test (SST) is the reference standard for diagnosing AI, with around 100,000 tests performed across the NHS annually. The SST is invasive, involving intravenous or intramuscular administration of Synacthen which stimulates the production of cortisol, the concentrations of which are then measured via blood taken before and at 30-60 minutes after administration. This requires a hospital outpatient or day attendance, with drug and hospital costs totalling £184.

A new home-based test is now available that uses a patient's saliva to measure the concentration of cortisone, which is a surrogate marker for free cortisol. This requires a testing pack to be posted to patients that contains simple sampling instructions, a swab to chew, a soft plastic container and a stamped-addressed envelope to return the pack directly to a laboratory. Cortisol is naturally secreted following a circadian rhythm, and so the saliva needs to be collected at a standardised time of day; a waking salivary cortisone (WSC) was chosen as the most practical way of standardising timings, as done at home with less confounding factors such as the stress of attending hospital influencing the result (Debono et al, 2009). This should provide us with a more physiological value that is close to peak cortisol. As well as being more convenient for patients, the test only costs £19 per patient.

A recent diagnostic accuracy study has compared SST and WSC (Debono et al, 2023). When using cutoffs of 95% sensitivity and specificity the study produced negative and positive prediction values for AI of 96% (95% CI; 90 to 99) and 95% (95% CI; 87 to 99), respectively. However, despite the potential costs saving of £165 per patient if WSC was to directly replace SST, this is not considered realistic given the slightly inferior diagnostic performance of WSC relative to the SST. Consequently, it is more likely to be adopted as part of a two-stage diagnostic strategy, with WSC used to screen for AI and the hospital-based SST used to test for AI in patients with an equivocal WSC result.

In this study, we report an economic evaluation of the two-stage diagnostic strategy described above, compared to the current diagnostic approach of using the SST for all patients suspected of having AI. The diagnostic thresholds chosen for WSC test were below 7 nmol/l to indicate AI and 17 nmol/l or higher to indicate the absence of AI; all other valid readings were considered to be diagnostically ambiguous.

This is undertaken by formulating a decision tree to look at costs and diagnostic performance for the two approaches, which is populated from the aforementioned diagnostic accuracy study. As such, the form of economic evaluation is a cost-consequences

analysis, with a time frame bounded by the production of a non-equivocal biochemical diagnosis (from either the WSC or SST).

Methods

A decision analytic model was developed to describe the costs and outcomes associated with different testing strategies. The evaluation takes an NHS perspective.

Model structure

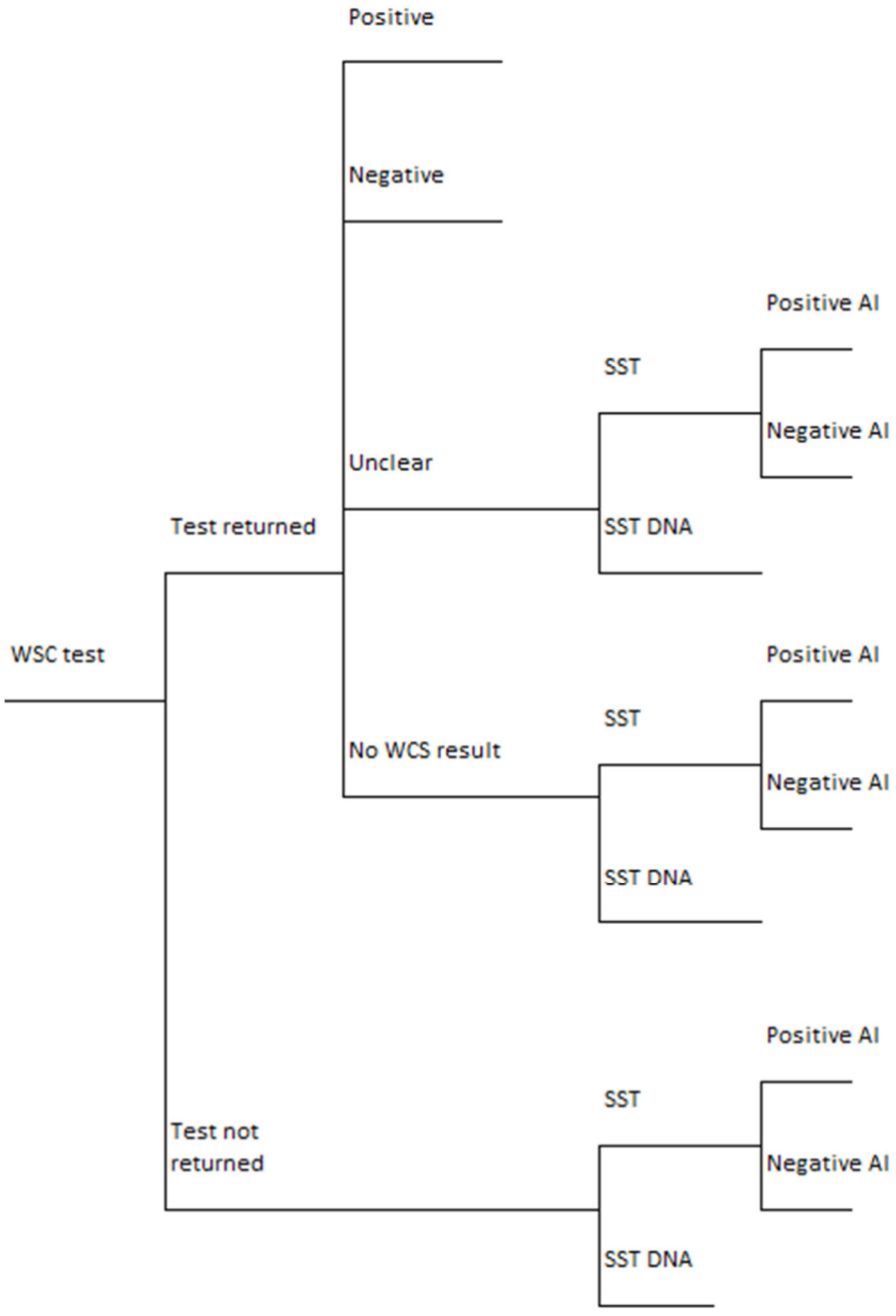
A decision tree describes a series of events by 'chance nodes' which are assigned probabilities for each eventuality associated with each event. The decision tree for the current diagnostic strategy is shown in Figure 1. In this, patients are referred for a SST, which they may or may not attend, then if they attend they are tested and are diagnosed as AI positive or AI negative.

Figure 1: Decision tree for the current strategy

Abbreviations:
SST= short synacthen test; AI= adrenal insufficiency; DNA= does not attend

The alternative diagnostic strategy that is assessed in our primary analysis is based on a patient being sent a WSC test, then if the diagnosis is unclear, they are referred for a SST at which point they follow the same route as in Figure 1. This is complicated slightly by the possibility that patients may not return their SCT, in which case it is assumed that they are then referred for a SST and then follow Figure 1. In addition, in some cases a WSC sample is returned that can not be adequately analysed, which in this strategy we assume will lead to a referral for a SST. In combination, this alternative diagnostic strategy is shown in Figure 2.

Figure 2: Decision tree for the alternative strategy



Abbreviations:
WSC= waking salivary cortisone; SST= short synacthen test; AI= adrenal insufficiency; DNA= does not attend

Model outcomes

Test performance is based on two sources of information. First, the WSC test performance is taken from the study by Debono and colleagues (2023). Second, the assumption is made that the SST is 100% correct; it only produces correct diagnoses. There is currently general support for this assumption, as SST is recognized as the biochemical gold standard. One further assumption is implicit within the parameterization of the model, specifically, that failure to return a WSC test, or give an adequate sample, or failure to attend a SST appointment is not related to the final diagnosis; as such, the underlying adrenal status of these patients is the same as that for the overall sample.

With these data and assumptions in place, we can calculate for each 'branch' the proportion of patients we expect to be true positive (TP), true negative (TN), false positive (FP), false negative (FN) or untested (i.e. do not attend (DNA).

Probabilities

The probabilities of events relating to the diagnostic outcomes are derived from the clinical analyses, with the clinical data for the primary analysis shown in Table 1.

Table 1: Clinical results associated with the primary economics analysis using waking salivary cortisone thresholds of 7 nmol/l and 17 nmol/l

	30 min Serum cortisol Immunoassay	
	AI	No AI
Positive test (WSC <7)	69	4
Negative test (WSC >= 17)	3	78
Unclear (7 <= WSC <17)	19	35
No WCS result	6	6

Abbreviations:
WSC= waking salivary cortisone; Ai= adrenal insufficiency

The probabilities for patients not returning the WSC test and not attending the SST, are based on expert opinion from using both tests within the NHS. Both are considered to be 3%.

Model costs

Test costs are based on estimates from Sheffield Teaching Hospitals (STH) and patient costs are based on patient questionnaire responses gathered within the Debono (2023) study. The patient questionnaires gathered responses from 157 (71%) patients. The data used to estimate costs associated with a SST were; method of transport to the hospital, expenditure

on transport (and parking if relevant), distance to and from the hospital (to generate a cost for car use) and time taken for attending the hospital (including travel). All unit costs are at 2020/21 price levels and are summarized in Table 2.

Table 2: Unit costs used in the model

Resource	Unit cost, £(2020/21)	Source
WSC test returned	18.51	Sheffield Teaching Hospitals
WSC test not returned	4.21	Sheffield Teaching Hospitals
SST test attendance	183.59	Sheffield Teaching Hospitals
SST DNA	0	Assumption
Patient costs associated with a SST	36.68	Patient questionnaire analysis
Patient costs associated with a WSC test	0	Assumption

Abbreviations:
WSC= waking salivary cortisone; SST= short synacthen test; DNA= does not attend

Analysis

The primary analysis is based on WSC thresholds of <7 nmol/l for positive diagnosis, >=17 nmol/l for negative diagnosis, and with all other values being considered equivocal. Scenario analyses were undertaken using different threshold values and taking a societal perspective that includes NHS, patient costs and associated production losses.

The alternative biochemical thresholds assessed are, (i) SST using Immunoassay, 400nmol cutoff, <6 &>=15 nmol/l, and (ii) SST using Immunoassay, 350nmol cutoff, <4 &>=10 nmol/l. The results relating to these thresholds when applied to the Debono (2023) patient cohort are given in the Appendix.

Results

The economic results for the primary analysis show that a two-stage diagnostic strategy costs £103 per patient less than the current diagnostic strategy (or 58% less). This is associated with a worse diagnostic performance with 7 out of 220 patients (3.2%) receiving a false diagnosis. This performance also includes a small increase in the number of diagnoses given, which is due to the greater engagement with the diagnostic process brought about by having two rounds of testing. The full breakdown of costs and outcomes is shown in Table 3.

Table 3: Costs and outcomes for the primary analysis

Diagnostic Strategy	Average cost per patient (£)	Outcomes				
		True positive	True negative	False positive	False negative	No diagnosis
Current	178.08	94	119	0	0	7
Two-stage	75.25	93	118	4	3	2
Difference	-102.84	-1	-1	4	3	-4

When the secondary analyses are considered, the use of alternative diagnostic cutoffs for SST has little impact on the results. The use of a societal perspective increases cost savings to £127 per patient (or a 59% reduction relative to current care).

Table 4: Differences in costs and outcomes relative to the current strategy for the primary and secondary analyses

Analysis*	Difference in average cost per patient (£)	Differences in outcomes				
		True positive	True negative	False positive	False negative	No diagnosis
Primary analysis	-102.84	-1	-1	4	3	-4
Scenario analysis 1	-104.41	-2	-3	6	4	-5
Scenario analysis 2	-114.61	-1	-2	6	3	-5
Scenario analysis 3	-127.00	-1	-1	4	3	-4

Notes: Scenario analysis 1 relates to SST using Immunoassay, 400nmol cutoff, <6 &>=15 thresholds (95% sensitivity and specificity). Scenarios analysis 2 relates to SST using Immunoassay, 350nmol cutoff, <4 &>=10 thresholds (95% sensitivity and specificity). Scenario analysis 3 relates to taking a societal perspective.

Discussion

We have used the results of the diagnostic study to evaluate two strategies for diagnosing adrenal insufficiency in patients referred to hospital for a SST to assess for a new diagnosis of adrenal insufficiency (or for recovery from a previous diagnosis of adrenal insufficiency. The current strategy involves referring all patients for a SST, whilst the alternative strategy is to send out a WSC test to all patients, with selective referral for a subsequent SST. The alternative strategy is shown to reduce NHS costs by 58% from £178.08 to £75.25. This reduces the diagnostic performance of the testing strategy in a small number of patients.

Changing the diagnostic cutoffs for SST had little impact on the results, whilst the inclusion of societal costs increased the costs of both strategies and the savings (to £127 per patient). One must take into consideration that this analysis is based on patients referred to hospital for a SST after they might have had an early morning serum cortisol which did not definitely exclude adrenal insufficiency. Waking saliva cortisone could also replace this early morning serum cortisol and save further costs. Cost savings generated by moving from a hospital based early morning cortisol blood screening test to a home based wakening salivary cortisone test needs to be looked at and factored into any future health economics. Recent NICE guidelines indicate even greater cost savings may be possible due to the use of higher current costs for SSTs (NICE 2024).

The development of a model for this analysis allows us to look at different diagnostic strategies that involve conditional sequential testing and considers inadequate samples, non-return of samples and non-attendance at hospital for SSTs. The added complexity of this analytical approach introduces some analytical uncertainties, 5 of which are of note.

First, the non-return and non-attendance rates are uncertain as the study does not reflect how the two strategies are implemented in routine clinical care. Consequently, we used expert opinion when estimating the non-return and non-attendance rates. Also, these two rates were assumed to be independent, when they are likely to be positively correlated (i.e. people who do not return a sample are also less likely to attend clinic). However, the impact of this on the overall results, and the differential between the two strategies, is expected to be small for rates that are considered plausible.

Second, the costs used in the analysis relate, principally, to Sheffield Teaching Hospitals. However, its costs are, on average, less than 5% different from national average costs (with a Market Forces Factor of 0.95231 in 2019/20 as reported in National Reference Costs). The costs for the WSC test are based on those of a single testing centre and would be expected to vary between centres and if test volumes were to increase significantly. However, these contribute only a small amount to the overall costs, which are dominated by hospital costs (including the costs of synacthen, which has a nationally negotiated price).

Thirdly, the societal costs, which include patient costs and lost productivity costs, are more uncertain due to a combination of patient recall (which will affect the accuracy of the estimates produced) and methodological uncertainties relating to the valuation of lost productivity (Krol and Brower, 2014). However, one thing is certain, and that is if a societal perspective is taken into account the size of cost reductions associated with the assessed screening strategies will increase.

Fourthly, it should also be noted that reduced costs and financial savings are not the same. The costs used within economic evaluation are generally based on the value of the underlying resources. In financial terms, reduced use does not always generate savings. Supplementary analyses would be required to estimate financial savings to the NHS (or an individual Trust) especially as NICE estimates for the cost of an SST are higher than when we carried out this analysis (NICE 2024).

Finally, one further uncertainty relating to the analysis comes about by the relatively short time horizon of the modelling. The use of a three-month time horizon limits the patient outcomes to those of the initial diagnostic label, however, the consequences of these are not explored in the modelling. Ideally, we would want to identify what happens to patients, including, what are the health and cost consequences of treatment (or in the case of false negative results, non-treatment). We surmise that the false diagnoses will have only small, short-term consequences as the 'mistake' produced by the two-stage strategy will be identified through re-presentation and subsequent SST (which would be clinically indicated, in preference to a WSC in such circumstances), or further clinical observations. However, further data would need to be collected to assess whether this is the case or not.

One further issue is worth further consideration. Our analysis is based on WSC thresholds associated with 95% sensitivity and specificity, but clearly, other thresholds relating to other levels of test performance can be defined. Higher levels of sensitivity and specificity will reduce the number of false diagnoses, but also reduce the cost savings. Unfortunately, larger sample sizes are required to undertake these analyses as the number of observed false diagnoses in our study are so small (n=7 for 95% sensitivity/specificity) and become fewer as sensitivity and specificity increase.

Four pieces of further research have been identified that would improve the accuracy and validity of our results:

1. Assessment of WSC thresholds associated with higher levels of sensitivity and specificity. This could come from a much larger study or pooling the results of multiple small studies (if statistically and clinically appropriate).
2. Identification of more robust non-return and non-attendance rate estimates.
3. Reproduction of our methods in other settings to assess their generalizability. This should cover other hospitals and patient populations, including paediatrics.
4. Assessment of the health and cost consequences of the two diagnostic strategies in terms of subsequent care.

When the importance of these four issues is also considered, in terms of their potential for altering the conclusions of our research or its widespread implementation, only the last two research topics are considered to be of significant value.

Conclusions

We have shown that a two-stage diagnostic strategy that involves a preliminary WSC test with subsequent SST contingent on the WSC results reduces NHS and societal costs. This strategy is also associated with a small increase in the number of false positive and negative diagnoses. The two-stage strategy is considered to be a valuable alternative to the current diagnostic process, due to its cost savings, that are associated with improved patient experience and avoid the need booking in a hospital appointment that can take several months. Other benefits include the enhanced resilience of the WSC testing process. Further

studies should look to build on our findings, by adopting the proposed two-stage diagnostic pathway in the routine clinical practice of other hospitals.

References

Rushworth RL, Torpy DJ, Falhammar H. Adrenal Crisis. *N Engl J Med*. Aug 29 2019;381(9):852-861.

Debono M, Elder CJ, Lewis J, et al. Home Waking Salivary Cortisone to Screen for Adrenal Insufficiency. *NEJM Evid*. Feb 2023;2(2):EVIDoa2200182.

Debono M, Caunt S, Elder C, et al. Real world evidence supports waking salivary cortisone as a screening test for adrenal insufficiency. *Clin Endocrinol (Oxf)*. Dec 2023;99(6):517-524.

Krol M, Brouwer W. How to estimate productivity costs in economic evaluations. Pharmacoeconomics. 2014 Apr;32(4):335-44. doi: 10.1007/s40273-014-0132-3. PMID: 24504850.

Debono M, Ghobadi C, Rostami-Hodjegan A, Huatan H, Campbell MJ, Newell-Price J, Darzy K, Merke DP, Arlt W, Ross RJ. Modified-release hydrocortisone to provide circadian cortisol profiles. J Clin Endocrinol Metab. 2009 May;94(5):1548-54.

NICE. Adrenal Insufficiency: Identification and Management Link. London: NICE, 2024. (Available at: https://www.nice.org.uk/guidance/ng243)

Appendix: Clinical results used in the secondary analyses

95% sensitivity/specificity, 400nmol cutoff

	30 min Serum cortisol Immunoassay	
	AI	No AI
Positive test (WSC <6)	61	6
Negative test (WSC >= 15)	4	85
Unclear (6 <= WSC <15)	19	33
No WCS result	5	7

95% sensitivity/specificity, 350nmol cutoff

	30 min Serum cortisol Immunoassay	
	AI	No AI
Positive test (WSC <4)	48	6
Negative test (WSC >= 10)	3	112
Unclear (4 <= WSC <10)	18	21
No WCS result	4	8