Serial measurement of the C-reactive protein is a poor predictor of treatment outcome in prosthetic joint infection

Philip Bejon1,2*, Ivor Byren1, Bridget L. Atkins1,3, Matthew Scarborough1, Andrew Woodhouse1, Peter McLardy-Smith1, Roger Gundle1 and Anthony R. Berendt1

1Bone Infection Unit, Nuffield Orthopaedic Centre, Windmill Road, Oxford OX3 7LD, UK; 2Nuffield Department of Medicine, Oxford University, John Radcliffe Hospital, Headley Way, Oxford OX3 9DU, UK; 3Department of Microbiology, John Radcliffe Hospital, Headley Way, Oxford OX3 9DU, UK

*Corresponding author. Bone Infection Unit, Nuffield Orthopaedic Centre, Windmill Road, Oxford OX3 7LD, UK. Tel: +44-1865-741841; Fax: +44-1865-738027; E-mail: pbejon@well.ox.ac.uk

Received 11 February 2011; returned 29 March 2011; revised 12 April 2011; accepted 13 April 2011

Objectives: Prosthetic joint infection is usually treated using surgery and antibiotics. The response to the treatment regimen is often evaluated using serial monitoring of plasma C-reactive protein (CRP) concentrations. In order to examine how useful this monitoring is, we calculated the sensitivity and specificity of CRP concentrations for predicting treatment failure.

Patients and methods: We examined 3732 CRP measurements from 260 patients who were treated by either two-stage revision or debridement and retention. We tested the association between CRP concentration and outcome using logistic regression models, and assessed sensitivity and specificity by using receiver operator curves.

Results: The areas under receiver operator curves for CRP concentrations predicting outcome ranged from 0.55 to 0.65.

Conclusions: CRP concentrations did not accurately predict treatment failure. Serial monitoring may not be of benefit.

Keywords: debridement and retention, biomarker, two-stage revision, sensitivity, specificity

Introduction

Arthroplasty is a highly cost-effective treatment for arthritis, but is complicated by a 1% incidence of prosthetic joint infection (PJI).1 Management of PJI most commonly includes surgery and prolonged antibiotic treatment. Unsuccessful treatment may lead to further surgery and ultimately to amputation.2

PJI may recur despite an apparent cure, and definitive diagnosis requires invasive sampling.2 Hence, it is attractive to use surrogate markers such as the C-reactive protein (CRP) in an attempt to predict outcome. Comprehensive meta-analysis supports using CRP to diagnose PJI at presentation,4 but smaller studies do not support using CRP to predict cure, at re-implantation during two-stage revision surgery for PJI.5–7

We analysed serial CRP measurements from two cohorts of patients with PJI including long-term clinical outcomes,4,8 in order to determine the clinical utility of CRP measurement during treatment of PJI.

Methods

As described previously, two clinical datasets were retrospectively gathered from patients with PJI managed in the Nuffield Orthopaedic Centre, Oxford, (i) by debridement, antibiotic and implant retention (i.e. ‘DAIR’) between 1 January 1998 and 30 April 2003 or (ii) by two-stage revision between 1 January 1999 and 30 April 2003.8,9

There was no research-related contact with patients, and informed consent was not required (as advised by our institutional review board). All activity was in accordance with the Declaration of Helsinki and national and institutional standards.

Case definition

Infection was defined as a clinical syndrome (persistent inflammation in the tissues around the implant, wound discharge or implant loosening) with any of the following: bacterial growth of an indistinguishable organism from two or more deep peri-prosthetic tissue samples; a neutrophilic infiltrate on histology of peri-prosthetic tissue; or a persistent sinus tract.3,8,9
Definition of treatment failure

Treatment failure was defined as (i) a draining sinus, (ii) further revision surgery (irrespective of the indication), or (iii) amputation of the affected limb. Undertaking further debridement with implant retention was a secondary outcome.

Antibiotic management

As described previously, intravenous antibiotic therapy was given for 6 weeks following surgery, irrespective of the CRP measurements. Patients undergoing DAIR were given prolonged oral antibiotics (mean duration 1.5 years).

Surgical management

Sinuses were excised and the material discarded. Intra-operative samples for culture and histology were taken as previously described. The integrity of the cement–bone or prosthesis–bone interfaces was tested. If the prosthesis was sound, DAIR was undertaken, with exchange of modular prosthesis components and removal of loose bone graft. If the prosthesis was loose, it was removed and the bone ends, medullary cavities and joint cavity debrided. Re-implantation was undertaken at a median of 120 days later.

CRP measurements

CRP measurements were taken weekly for the first 6 weeks, then at scheduled outpatient visits (typically three times during the first year then twice yearly), using a BioStat kit and an Aeroset analyser (Abbott, Maidenhead, UK). The limits of measurement are between 8 and 285 mg/L. CRP concentrations below or above the limits of measurement were substituted by values of 4 and 300 mg/L, respectively. The laboratory was accredited by the UK Clinical Pathology Accreditation scheme.

Analysis

STATA version 10 (Stata Corp., TX, USA) was used to fit multiple fractional polynomial regression models for CRP over time and clinical outcome, selected by the Royston and Altman algorithm; to plot receiver operator curves (ROCs); and to calculate sensitivity, specificity and likelihood ratios. We divided the CRP data into (i) the period of initial decline and (ii) the subsequent plateau. This resulted in four different datasets: (i) during the first 180 days after DAIR; (ii) following the first 180 days, using time to treatment failure or to the end of follow-up; (iii) during the first 120 days after the first stage of a two-stage revision (excluding CRPs taken after re-implantation); and (iv) after re-implantation, excluding the first 28 days post-re-implantation. CRP measurements taken during hospital admissions for inter-current illness were excluded.

Results

There were 2326 measurements of CRP from 109 patients undergoing DAIR (comprising 51 hip replacements, 50 knee replacements and 8 other joints; 84 of which were primary and 25 revised joint replacements) and 1406 separate measurements and 8 other joints; 84 of which were primary and 25 revised joint replacements) and 1406 separate measurements from 151 patients undergoing two-stage revision (comprising 71 hip, 76 knee and 4 elbow revisions; 62 of which were primary replacements and 89 of which were previously revised). The median lengths of follow-up with CRP measurements were 1.5 years following DAIR (interquartile range (IQR) 1.1–2.2 years) and 2 years following two-stage revision (IQR 4 months to 5 years). The median number of CRP measurements was 10 (IQR 5–15) after DAIR and 7 (IQR 3–13) after two-stage revision. Examples of the CRP concentrations of individual patients over time are shown in Figure S1 (available as Supplementary data at JAC Online).

DAIR

The profiles of falling CRP concentrations during the first 180 days after DAIR are shown in Figure 1(a). After adjusting for time, the ratio of CRP readings for treatment failure was 1.57 [95% confidence interval (CI) 1.07–2.0, P=0.02]. In other words, CRP readings were, on average, 57% higher in patients experiencing treatment failure.

Since time since surgery was more strongly associated with CRP concentration than treatment outcome, we were unable to produce ROCs using absolute CRP concentrations. Instead, we produced ROCs using the relative CRP level after adjusting for time. The area under the ROC (AUROC) was 0.6, indicating that CRP was a poor test (Figure S2, available as Supplementary data at JAC Online). For example, if a CRP concentration of >1.4 times the concentration expected (based on the decline seen during non-failing treatment) was considered to indicate a ‘positive result’, then the specificity was 80% for predicting failure and a sensitivity was 22%, with positive and negative likelihood ratios of 1.1 and 1.03, respectively. We were unable to select any CRP threshold that gave reasonable sensitivity and specificity (Figure S2).

CRP did not predict the secondary outcomes of early failure (i.e. for treatment failure identified during the first 180 days, n=6, the CRP ratio was 0.84, 95% CI 0.6–1.14, P=0.3) or additional debridement with implant salvage (1.1, 95% CI 0.97–1.24, P=0.12, n=24).

CRP fell during the first 120 days after surgery. The CRP profile was not significantly different in patients experiencing treatment failure (the ratio of current to previous month’s CRP was 1.035, 95% CI 0.99–1.07), but a significant increase over time in subjects who did experience failure (a ratio of 1.15/month, 95% CI 1.06–1.25, P=0.006 for the interaction between time and treatment failure (Figure 1b). When CRP was used as a test to predict failure within 1 year, the AUROC was 0.65, indicating marginal usefulness. For instance, a CRP of >40 mg/L gave a specificity of 80% and a sensitivity of 50%, with positive and negative likelihood ratios of 2.5 and 1.6, respectively.

Two-stage revisions

CRP fell during the first 120 days after surgery. The CRP profile was not significantly different in patients experiencing treatment failure (ratio=0.7, 95% CI 0.5–0.91, P=0.064) (Figure 1c). CRP did not predict additional debridement between stages (ratio=1.23, 95% CI 0.8–1.85, P=0.3), but did predict delayed re-implantation (ratio=1.03 for each month of delay, 95% CI 1.01–1.05, P=0.003).

CRP at the time of re-implantation was not associated with treatment failure (ratio=0.84, 95% CI 0.65–1.09, P=0.24), or with positive microbiology at re-implantation (ratio=0.9, 95% CI 0.73–1.1, P=0.34, n=21).

After re-implantation, there was no significant change in CRP over time among patients whose treatment was successful (ratio=0.99/month, 95% CI=0.98–1.03), but an increase in...
patients who experienced treatment failure (ratio=1.02/month, 95% CI 0.97–1.07, P=0.038 for the interaction between time and treatment failure) (Figure 1d). However, CRP was not a useful test to predict failure within 1 year (AUROC=0.55).

We also examined the absolute neutrophil counts. The neutrophil count fell over time after DAIR ($-2.6\times10^9$ cells/L/week, $P<0.0005$) and after two-stage revision ($-1.2\times10^9$ cells/L/week, $P<0.0005$), but was not associated with treatment failure after DAIR ($0.23\times10^9$ cells/L higher, 95% CI $-1.3$ to $1.8$, $P=0.8$) or after two-stage revision ($0.38\times10^9$ cells/L higher, 95% CI $-0.4$ to $0.8$, $P=0.6$), and so was not analysed further.

**Discussion**

Treatment failure after DAIR was associated with a CRP that was slow to normalize post-operatively, or a high CRP during long-term follow-up. Treatment failure after two-stage revision was associated with a high CRP during long-term follow-up, but not with the rate of post-operative normalization. However, CRP could not be recommended as a diagnostic test based on the sensitivity and specificity values indicated by ROCs. This does not reflect limited power of the study, but the wide scatter of individual readings in both outcome groups, as found in previous studies.5–7 We did not detect statistically significant trends over time in the white cell count, neutrophil count or platelet count (data not shown), and did not measure procalcitonin or interleukin-6.

Measuring a single CRP may cost 15 US dollars (i.e. $56000 for the 3732 tests analysed here or $215/patient), but inappropriate management decisions (for instance, invasive sampling or delayed re-implantation) may be much more costly.

Observational data are prone to bias. For instance, high CRP during two-stage revision was associated with delayed re-implantation. Since CRP was not, in fact, associated with treatment failure or additional surgical debridement, this indicates that implantation was delayed simply by the clinician’s response to the high CRP. However, irrespective of potential biases, the very wide scatter of CRP readings irrespective of outcome indicates the limitation of the test.

Serial CRP measurements are cheap, biologically plausible, predict the response to treatment for endocarditis10 and so are readily included in the care bundles for infectious diseases. During treatment for PJL, we found that CRP monitoring was a poor test of cure. In order to avoid triggering needless
interventions on the one hand, or delaying clinically indicated interventions on the other, we recommend against routinely monitoring CRP during treatment of PJI.

**Funding**
The study was carried out as part of a routine service audit. P. B. is supported by the NIHR Biomedical Research Centre in Oxford.

**Transparency declarations**
I. B. has received honoraria for serving on advisory boards (Pfizer) and lecture fees (Pfizer and Nordic Pharma). P. M.-S. is advisor to Wright Medical Technologies and receives royalties from the Corin Group. A. R. B. has received honoraria for serving on advisory boards (Pfizer and Macrochem), for serving on speakers bureaus (Merck) and for producing sponsored non-promotional educational materials (Merck). All other authors: none to declare.

**Supplementary data**
Figure S1 and S2 are available as Supplementary data at JAC Online (http://jac.oxfordjournals.org/).

**References**