6-thioguanine nucleotide monitoring in azathioprine and mercaptopurine monotherapy for the treatment of inflammatory bowel disease (Protocol)

Mateen BA, Patel M, Akobeng AK, Gordon M, Hayee B

6-thioguanine nucleotide monitoring in azathioprine and mercaptopurine monotherapy for the treatment of inflammatory bowel disease (Protocol).

Cochrane Database of Systematic Reviews 2021, Issue 10. Art. No.: CD014795.
DOI: 10.1002/14651858.CD014795.

www.cochranelibrary.com
6-thioguanine nucleotide monitoring in azathioprine and mercaptopurine monotherapy for the treatment of inflammatory bowel disease

Bilal Akhter Mateen1,2, Mehul Patel2, Anthony K Akobeng3, Morris Gordon4, Bu'Hussain Hayee1,2

1Department of Gastroenterology, Kings College Hospital NHS Foundation Trust, London, UK. 2School of Life Sciences & Medicine, Kings College London, London, UK. 3Pediatric Gastroenterology, Sidra Medicine, Doha, Qatar. 4School of Medicine, University of Central Lancashire, Preston, UK

Contact address: Bilal Akhter Mateen, bilal.mateen@nhs.net.

Editorial group: Cochrane Gut Group.


Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

ABSTRACT

Objectives
This is a protocol for a Cochrane Review (intervention). The objectives are as follows:

To determine the efficacy and safety of a 6-thioguanine nucleotide (6-TGN) metabolite-based dosing strategy for maintenance of remission using azathioprine or mercaptopurine in Crohn’s disease and ulcerative colitis.
BACKGROUND

Inflammatory bowel disease (IBD) is a non-infectious chronic inflammatory disease of the gastrointestinal (GI) tract. It is commonly subdivided into two pathognomonic patterns of inflammation, where Crohn’s disease can affect any segment of the GI tract from mouth to anus, and ulcerative colitis is limited to the colonic mucosa (Sairenji 2017). Approximately 7 million people have IBD globally, leading to over 1 million years lost to disability annually (Alatab 2020).

Description of the condition

Treatment of IBD is split into two major therapeutic aims: 1) inducing remission and 2) maintaining remission. An array of different induction agents have been demonstrated to be effective in Crohn’s disease, Akobeng 2003; Chande 2015; Rezaie 2015; Lim 2015; MacDonald 2016; Nelson 2018; Abbass 2019, and ulcerative colitis (Lawson 2006; Baumgart 2008; Marshall 2010; Bickston 2014; Rosenfeld 2015; Sherlock 2015; Wang 2016b). Similarly, there are several medical therapies for maintenance of remission in Crohn’s disease, Kuenzig 2014; Akobeng 2016; Chande 2016; Battat 2017; Townsend 2020, and ulcerative colitis (Marshall 2012; Wang 2015; Davies 2020; Timmer 2016; Wang 2016a). In this review we will concern ourselves only with the latter objective of maintenance and the use of one specific family of therapeutic agents, the thiopurines, which includes mercaptopurine and its pro-drug azathioprine.

Description of the intervention

Over 60 years since the immunosuppressive properties of azathioprine and mercaptopurine were first demonstrated (Schwartz 1958; Calne 1962), both remain commonly utilised therapeutic agents in the medical management of IBD (Harbord 2017; Lamb 2019; Torres 2019). Previous systematic reviews and meta-analyses assessing the efficacy of azathioprine have demonstrated that doses between 1.0 mg/kg/day and 2.5 mg/kg/day are more effective than placebo for maintenance of remission (Chande 2016; Timmer 2016). However, weight-based dosing is complicated by the fact that there is significant inter-individual variation in the metabolism of thiopurine agents, due to allelic polymorphism in the thiopurine-S-methyltransferase (TPMT) enzyme, which is principally responsible for the metabolism of azathioprine into its pharmacologically active form (Weinshilboum 1980; Vuchetich 1995; Relling 2011).

The clinical relevance of increased TMPT activity is that patients may be receiving subtherapeutic doses when a generic weight-based dosing protocol is utilised, Cuffari 2000; Cuffari 2006, due to preferential conversion of azathioprine into IBD-inert metabolites such as methylmercaptopurine (Wright 2004). Moreover, there is also an increased risk of hepatotoxicity in such circumstances (Dubinsky 2000), whereas in individuals with reduced TMPT activity, there is an increased risk of other potentially life-threatening adverse events (e.g. myelotoxicity) due to accumulation of the active metabolite 6-thioguanine nucleotide (6-TGN) (Cuffari 2006; Hindorf 2006). Approximately 10% to 20% of individuals cease azathioprine treatment due to side effects, hepatotoxicity, or dose-limiting intolerance (Pearson 1998; Gemollon 2008).

More recently, the enzyme Nudix hydrolase 15 (NUDT15) has also been identified as a genetic determinant of thiopurine-related haemotoxicity (Moriyama 2016). Specifically, loss-of-function germline variants lead to the accumulation of the active thiopurine metabolite thioguanosine triphosphate (TGTP), which results in haemotoxicity (Cargin 2018). The active metabolite TGTP would normally be inactivated by conversion to the monophosphate by-product due to the nucleotide diphosphate activity of NUDT15 (Dean 2020). The relevant NUDT15 variants are thought to be most prevalent in Asian populations, but more recent work has demonstrated their contribution to thiopurine-related toxicity in South American and European populations as well (Soler 2018; Schaeffeler 2019, Walker 2019). In essence, this is another key genetic variant that may contribute to the success or failure of thiopurine weight-based dosing strategies.

How the intervention might work

The thiopurine family of therapeutic agents, including azathioprine and mercaptopurine, are metabolised into purine analogues. The simplified metabolic pathway from azathioprine to 6-TGN is illustrated in Figure 1. The full proposed pathway is described elsewhere (Lennard 1992). The mode of action of 6-TGN is not fully understood. One potential mechanism is that they act as antimetabolites, specifically exerting their influence by inducing post-replication DNA mismatch repair-related cytotoxicity in T-cell by acting as competitive antagonists for the naturally occurring DNA nucleobase guanine (Tidd 1974; Karran 2007). Alternatively, 6-TGN and its phosphorylated products have been shown to inhibit several pro-inflammatory pathways and gut-homing mechanisms via modulation of Rac1 GTPase. This may also explain the well-established usefulness of the thiopurines in IBD (Tiede 2003; Marinkovic 2014a; Marinkovic 2014b; Seinen 2016). In either circumstance measuring 6-TGN concentrations is thought to provide direct insight into the quantity of the presumed therapeutic substrate, thereby accounting for individual variation in metabolism, instead of the more adverse-event prone circumstance of relying on weight and TMPT genotype for the dosing (Colombel 2000). Whilst the primary site of action is leukocytes, due to ease of access and the simpler biochemical processes involved (Coulthard 2016), the 6-TGN assays utilised in clinical practice are based on peripheral red cell concentrations as a proxy (Dooley 1982; Lennard 1983; Hofmann 2012; Vikingson 2013). The literature suggests that an “efficacious” erythrocyte 6-TGN concentration is anywhere between 235 and 450 pmol/L for the method described by Lennard (Osterman 2006; Moreau 2014; Estevinho 2017; Warner 2018). There are several other assay methods for 6-TGN quantification which have different thresholds for efficacious concentrations (e.g. Dervieux 1998).
Figure 1. Metabolism of Azathioprine: A simplified illustration of the breakdown of azathioprine into its active and inactive metabolites.

Metabolism of Azathioprine

Legend: Azathioprine is principally metabolised in the liver from its pro-drug form (azathioprine) to the active metabolite 6-mercaptopurine (6-MP), which is subsequently converted by the thiopeurine methyltransferase (TMPT) enzyme into 6-TGN. 6-TGN is then phosphorylated into metabolites which act as competitive substrates in purine-derived nucleobase synthesis, and once incorporated into DNA has downstream immunomodulatory effects.

Why it is important to do this review

Clinical guidelines highlighting the potential usefulness of a metabolite-based dosing strategy have tended to draw on a large body of observational research which suggests that reaching the aforementioned thresholds is associated with a greater propensity to have maintained remission at medium-term (e.g. 3 to 12 months) time points (Osterman 2006; Moreau 2014; Estevinho 2017). However, this potential prognostic property of 6-TGN is not the same as evidence of effectiveness for its use as a dose-modification strategy. In fact, there appears to be a paucity of randomised controlled trial-based evidence in this area. We are only aware of two randomised controlled trials addressing the effectiveness of a 6-TGN dose-modification strategy for azathioprine (Reinshagen 2007; Dassopoulos 2014). The first randomised controlled trial was stopped due to inadequate recruitment. Still, the authors argued that the available data suggested improved outcomes associated with a metabolite-based dose-adjustment strategy (Dassopoulos 2014). In contrast, the second study concluded that there was no benefit compared to weight-based dosing (Reinshagen 2007). These conflicting results have likely contributed to the discordance amongst national and international consensus groups; whilst some of these groups have refrained from making a recommendation on the topic of metabolite monitoring-based approaches for azathioprine therapy in IBD (Harbord 2017; Torres 2019), others recommend its use whilst acknowledging that the justifying evidence is of poor quality (Benkov 2013; Feuerstein 2017; Lamb 2007; Dassopoulos 2014).
OBJECTIVES

To determine the efficacy and safety of a 6-thioguanine nucleotide (6-TGN) metabolite-based dosing strategy for maintenance of remission using azathioprine or mercaptopurine in Crohn’s disease and ulcerative colitis.

METHODS

Criteria for considering studies for this review

Types of studies

We will consider all randomised active-comparator controlled trials of any duration for inclusion. We do not intend to include cluster randomised or crossover trials, but will record if any are identified in the course of the review.

Types of participants

We will consider all participants in whom azathioprine or mercaptopurine was used to treat Crohn’s disease or ulcerative colitis, with or without a preceding period of induction of remission, for inclusion in the review. We will apply no age restrictions.

For studies without an induction period (i.e. in which recruited participants were already in remission), the definition of remission based on conventional composite clinician-assessed and patient-reported tools or endoscopic criteria will be as follows (see Peyrin-Biroulet 2016 for a review).

• For individuals with Crohn’s disease, the clinical criteria will be a Crohn’s Disease Activity (CDAI) < 150 points (Best 1976; Thia 2011), or Harvey-Bradshaw Index score ≤ 4 (Harvey 1988; Vermeire 2010), and complete discontinuation of corticosteroids, irrespective of concomitant use of other prophylactic medication. Endoscopic criteria will be a Simple Endoscopic Score for Crohn’s Disease (SES-CD) score ≤ 2 (Daperno 2004).

• For individuals with ulcerative colitis, the clinical criteria for remission will be a Simple Clinical Colitis Activity Index (SCCAI) score ≤ 4 (Walmsey 1998), or a score of ≤ 1 on the patient-reported assessment of disease activity subsections of the Mayo score (Schoeder 1987), and complete discontinuation of corticosteroids, irrespective of concomitant use of other prophylactic medication. Endoscopic criteria for remission will be a Mayo endoscopy subscore score of ≤ 1 (Schoeder 1987), or an Ulcerative Colitis Endoscopic Index of Severity (UCEIS) score of ≤ 1 (Travis 2012), or a modified-Baron score of ≤ 1 (Feagan 2005; Paine 2014).

• For studies with an induction period, the definition used for active disease when recruiting individuals with Crohn’s disease or ulcerative colitis will be defined by the failure to meet the aforementioned composite clinician-assessed and patient-reported tool-based or endoscopic criteria-specific thresholds for remission.

Types of interventions

Trials of oral azathioprine or mercaptopurine therapy where therapeutic dose was defined by achieving and maintaining a specific (erythrocyte 6-TGN) metabolite threshold. Active comparators will include a weight-based dosing strategy or any other biomarker-defined therapeutic dosing strategy.

There are other important metabolites that are often measured, such as methylmercaptopurine (MeMP) and derived variables such as the MeMP:6-TGN ratio that are more often used for prognosticating the risk of adverse effects and determining whether treatment with azathioprine and allopurinol combination therapy might be more appropriate than thiopurine monotherapy, respectively. Both of these are outside the scope of this review.

Types of outcome measures

We will assess outcomes based on an intention-to-treat approach. For studies with an induction phase preceding the maintenance phase, the expectation is that studies will re-randomise individuals at the start of the latter as is standard practice in induction + maintenance trials, which will permit comparison of just the maintenance phases. All studies that fail to re-randomise will be deemed controlled (non-randomised) prospective maintenance studies, and will thus be excluded from the primary analysis.

Primary outcomes

• The proportion of participants that maintained clinical or endoscopic remission, up to and including any follow-up point that is more than six months from a) recruitment into a maintenance trial, or b) from the end of the induction phase. It is important to note that definitions of failure to maintain remission do not always conform to the aforementioned criteria, therefore we will explicitly extract the specific definition used in each study. Where appropriate, we will perform sensitivity analysis to compare variation in this outcome based on different definitions of remission.

• The overall safety of the treatment strategy, as part of the reason for weight-based dosing is to reduce the likelihood of adverse effect-related treatment cessation, which is a significant cause of treatment failure. We will extract data with the aim of exploring the toxicity of different dosing strategies defined by the occurrence of an adverse event. The primary outcomes will be the overall rate of adverse events, and the overall rate of serious adverse events, for which we will extract data as defined by each study. To provide greater interpretability of how comparable different studies’ definitions are, we will also extract data broken down into the following groups: opportunistic infections, pancreatitis, haematological complications (i.e. agranulocytosis, aplastic anaemia, etc.), hepatic complications (hepatotoxicity, including elevation of serum alkaline phosphatase, bilirubin, and/or serum transaminases), other systemic manifestations (e.g. nausea, lethargy, etc.), cancer, or death.

Secondary outcomes

• Short-term remission: the proportion of participants who demonstrate evidence of clinical or endoscopic remission, up to and including any follow-up point that is less than six months from recruitment, or the end of the induction phase.
• Effective dose: we will extract data to compare the dose required in the metabolite-based dosing strategy compared to the weight-based dosing strategy, to determine the absolute magnitude of the difference between doses in those with sustained remission at each predefined study endpoint.

• Quality of life: where relevant data are reported, they will be extracted to determine if there are dosing strategy-specific differences in quality of life at any of the predefined study time points.

We will additionally consider a separate (per-protocol) analysis of studies with an induction phase, assuming a sufficient number of controlled (non-randomised) prospective maintenance studies are identified. The purpose of this analysis is to determine the likelihood of sustained remission in the context of different dosing strategies for azathioprine or mercaptopurine, in those individuals that successfully navigated the induction phase and thus demonstrated tolerance to side effects.

Search methods for identification of studies
The scope, processes, and resourcing of the search methods have been specifically selected to conform to the guidance in Chapter 4 of the Cochrane Handbook for Systematic Reviews of Interventions (Higgins 2021).

Electronic searches
With the support of the Cochrane Gut information Specialist, an electronic search strategy has been created using a combination of MeSH (Medical Subject Headings) terms and text phrases related to IBD, thiopurines, and the 6-TGN metabolite. The electronic search will comprise the Cochrane Central Register of Controlled Trials (CENTRAL) (the Cochrane Library) (Appendix 1); MEDLINE (Ovid) 1966 to present (Appendix 2); and Embase (Ovid) 1980 to present (Appendix 3), to identify all relevant published studies.

Searching other resources
We will handsearch conference proceedings and clinical guideline references to identify additional studies. Specifically, we will screen all abstracts published in a peer-reviewed journal supplement from one of the following conferences: United European Gastroenterology (UEG), British Society for Gastroenterology (BSG), European Crohn’s and Colitis Organisation (ECCO), Crohn’s Colitis Congress, and Digestive Disease Week (DDW), from the earliest online record or the year 2000 (whichever is later), to present day. Moreover, we will screen the reference lists of the most recently published guidance document pertaining to the management of patients with IBD, published by the United European Gastroenterology (UEG), the British Society for Gastroenterology (BSG), and the American Society for Gastroenterology (ASG).

Data collection and analysis
The data collection, extraction, and analytic processes have been selected to conform to the guidance in Chapters 5 to 10 of the Cochrane Handbook for Systematic Reviews of Interventions (Higgins 2021)

Selection of studies
We will initially screen trials identified by the search strategy based on title and abstract to determine their broad relevance to the current review and thus prevent introducing bias resulting from incomplete reporting (i.e. studies only published as conference proceedings). Two review authors (BAM and MP) will undertake this process in an independent, unblinded manner. Studies will be designated to one of three categories: 1) ineligible, 2) require more information, or 3) eligible. In cases where both review authors agree that the abstract is ineligible, it will be discarded. Where there is disagreement between review authors, the study will be assigned to the ‘require more information’ group without discussion. We will undertake a final round of full-text screening for the group in which title and abstract screening was deemed insufficient. We will list full texts excluded at this stage and the reasons for their exclusion in the ‘Characteristics of excluded studies’ table. One of the non-screening authors will arbitrate any persistent disagreement, and the final consensus decision will be recorded. All retained studies will progress to the data extraction phase. We will display the selection process in a flow chart as recommended in the PRISMA statement (Moher 2009), illustrating the total numbers of retrieved references and the numbers of included and excluded studies.

Data extraction and management
Two review authors (BAM and MP) will independently extract data using a specially designed data collection form consisting of the following items.

• Trial design: sequence generation, blinding of participants and clinicians, active comparator, losses to follow-up (and reasons); for studies with an induction phase, the design of both the initial randomisation at recruitment and re-randomisation at the initiation of the maintenance phase will be extracted.

• Recruitment and follow-up: inclusion criteria, exclusion criteria, prespecified follow-up dates (specifically including the date at which participants were deemed treatment failures for not achieving remission in studies with an induction arm).

• Treatment strategy: any of the following actions in response to a 6-TGN result are deemed eligible: split dosing, increase dose, or decrease dose. Initiation of allopurinol as part of a combination (azathioprine and allopurinol) strategy will be excluded.

• Sample characteristics: ethnicity (counts and proportions), age (mean and standard deviation, or median and range), sex assigned at birth (counts and proportions), duration of disease (mean and standard deviation, or median and range), type of IBD (counts and proportions), distribution of disease (counts and proportions, based on the Montreal Classification (Satsangi 2006)), severity of disease at recruitment (counts and proportions, based on the aforementioned severity scoring tools), concomitant medications administered (counts and proportions), and genetic polymorphisms/phenotype (count and proportions for polymorphisms or the mean and standard deviation for phenotype activity, or both).

• Outcomes: all information pertaining to the primary and secondary outcomes specified in the Types of outcome measures section of the protocol.

Any disagreements between the two review authors (BAM and MP) will be resolved by discussion or through arbitration with a third review author where necessary, and the final consensus decision recorded. We will tabulate results based on several approaches, namely intention-to-treat for all studies, and additionally per-protocol for studies with an induction arm resulting in treatment failures to induction being excluded.
Assessment of risk of bias in included studies

Two review authors (BAM and MP) will assess the methodological quality of the included studies using Cochrane's risk of bias tool (described below) (Higgins 2011). Assessment will be based on examining the following risk of bias domains using a three-point ordinal scale: ‘low risk of bias’, ‘unclear’ (uncertain risk of bias), or ‘high risk of bias’, as follows.

Allocation of intervention

We will assess four aspects of the allocation process: sequence generation, concealment of allocation, blinding of participants, and blinding of the treating physicians. We will score the risk of bias for this domain as follows:

- high risk of bias: absence of a clear description of any two aspects or a weak methodological approach to two or more aspects;
- unclear risk of bias: if only one aspect is either methodologically flawed or absent;
- low risk of bias: if all four of the aforementioned aspects are present and deemed appropriate.

For studies with an induction phase, we will carry out the assessment of allocation of intervention primarily with regard to the maintenance phase re-randomisation.

Outcome measurement

Detection bias may be introduced based on the subjectivity of the outcome measurement criteria used to define remission and maintenance of remission.

In cases where a clinician-rated or a patient-rated measurement tool is used, and neither participant nor clinician is blinded to the intervention, we will make a judgement of high risk of bias. Where blinding is part of the study protocol, this will be judged unclear risk of bias; or similarly, even when blinding is not used, a combined patient- and clinician-rated tool may mitigate some biasing effect of knowing intervention-related information, and will thus also be judged as unclear risk of bias. Finally, the use of objective markers such as endoscopic or histological evidence of remission to define study endpoints will be judged as low risk of bias. Moreover, where an independent (i.e. not involved in care provision) observer-rated tool is utilised, this will also be judged low risk of bias.

Incomplete outcome data

Due to the risk of attrition bias, any study with more than 20% of initially recruited individuals lost to follow-up will be judged as high risk of bias. The method for addressing this risk as part of a sensitivity analysis is described in Sensitivity analysis. We will deem studies with between 10% and 20% loss to follow-up as unclear risk of bias. If there is less than 10% loss to follow-up, or statistical evidence of missingness completely at random, the study will be judged as at low risk of bias. Proportions instead of absolute counts are utilised to define the thresholds to penalise smaller sample size studies in which loss to follow-up would likely have greater impact on the results.

Selective outcome reporting

We will assess all studies referencing a prospectively published or registered protocol as at low risk of bias; otherwise, we will categorise studies as unclear risk of bias. We will then downgrade these designations in a step-wise fashion if there is evidence of the following issues in the trials:

- results reported in narrative format (i.e. not significant), but the numerical estimates are not provided;
- significance is assessed for either changes in score or final absolute score, but not both; and
- only a subset of time points at which the outcome was assessed are reported.

Other sources of bias

One example of a source of bias not previously mentioned is that of deviation from the intended intervention, that is the potential lack of explicit guidance in the protocol for the case in which a participant is deemed to have met clinical criteria for remission, but has not achieved the predefined metabolite profile based on 6-TGN concentrations. We will assess studies that fail to address this at the protocol stage as at unclear risk of other bias, as physicians may have deviated from the intervention and terminated up-titration of the thiopurine agent early based on clinical response and not metabolite profile. We will assess studies that provide clear instructions in the protocol on how this should be handled as at low risk of bias.

Other potential sources of bias requiring assessment and the formulation of definitions for each of the aforementioned risk of bias domains may be identified in the course of data extraction. Any retrospectively added criteria will be explicitly noted and their impact on the interpretation of the results discussed.

Study bias summary score

Following assessment of the above risk of bias domains, we will assign each study a summary rating of low risk, high risk, or unclear risk of bias based on all available evidence in the publication. We will use the following definitions to determine the summary outcome.

- High risk of bias: one or more risk of bias domains assessed as high risk of bias.
- Unclear risk of bias: no risk of bias domains assessed as high risk, and more than half assessed as unclear risk.
- Low risk of bias: no risk of bias domains assessed as high risk, and more than half assessed as low risk.

Any disagreements in score for an individual bias domain or for the summary score will be discussed between the two review authors performing the risk of bias assessment; if consensus cannot be reached, a third review author will act as arbiter, and the majority decision will be recorded.

Measures of treatment effect

The primary outcomes are dichotomous, and as such will be presented as risk ratios (RR). Where the secondary outcomes are of a different form (e.g. continuous/ordinal values from quality of life scales), we will report the mean difference (MD), or the standardised mean difference (SMD) if different scales are used.

We will present all reported outcome data with a 95% confidence interval (CI), and P values for the significance of the point-estimate relative to the null hypothesis of no difference between groups.
Where the pooled outcome data and the resulting risk ratio confidence interval do not straddle the position of the null effect (e.g. an RR of 1), we will additionally calculate the number needed to treat for an additional beneficial outcome (NNTB), and the associated 95% CI.

**Unit of analysis issues**

Ideally, all analysis would be undertaken for each thiopurine agent (i.e. azathioprine and mercaptopurine) separately, as there is well-documented variation in metabolism of the pro-drug (azathioprine) to the active compound (mercaptopurine). However, it is common for studies describing thiopurine-based interventions to apply a crude conversion factor of 2.5x to convert a patient’s mercaptopurine dose into an azathioprine equivalent dose when reporting sample characteristics. Moreover, even when these subgroups are described separately, the final outcome is often provided for the entire sample, without stratification by specific thiopurine. If these issues are present in the final review, then we will use the aforementioned crude conversion factor to generate a consistent expression of doses across studies (i.e. azathioprine-equivalent doses). However, we will also analyse the results by specific subgroup where possible, to determine whether the results are consistent across both the pro-drug (azathioprine) and the active agent (mercaptopurine).

**Dealing with missing data**

For all trials published only in abstract format for which the full text is not retrievable, or where additional information is needed for the purposes of the review, we will contact the corresponding author to determine whether a protocol, draft manuscript, or another record of the results can be sourced. Where no alternate source is available, we will assess the available abstract or text for suitability; the sole criterion for inclusion will be if sufficient information pertaining to one of our primary outcomes is reported, in which case the study will be retained (with sensitivity analysis to ascertain the impact of its inclusion/exclusion); otherwise, the study will be excluded and a record made to reflect this decision. For included studies with less than 80% of the total number of participants with a recorded outcome at follow-up, assuming this was not adequately addressed by contacting the authors, we will carry out an additional sensitivity analysis (see Sensitivity analysis).

**Assessment of heterogeneity**

We will determine clinical heterogeneity based on differences in the definition of the disease state (i.e. the criteria for diagnosing remission or active disease) of the sample at the relevant time points. We will not pool results where the characteristics are deemed (by consensus) to be uniquely irrelevant. Where clinical heterogeneity does not contraindicate pooling, we will then statistically assess the results for the presence of heterogeneity using the Chi^2 test. To accommodate the low power of the Chi^2 test in the absence of a large pool of studies and when studies have small sample sizes - both of which we expect to be the case in this review - the P value for significance will be set at 0.1. This approach will be complemented by the use of the I^2 statistic to quantify the degree of heterogeneity. This measure estimates the proportion of the total variation across studies that can be ascribed to heterogeneity rather than sampling error (i.e. chance). Overlapping bounds for the interpretation of the I^2 statistic are commonly utilised, which broadly correspond to 30% and above indicating moderate heterogeneity, and 50% and above indicating substantial heterogeneity. We will not restrict pooling by an a priori-defined tolerable upper bound on the I^2 statistic, but rather use it as part of a multipart assessment (including the risk of bias assessment as discussed above) to determine if reporting of the results is informative and appropriate. Where we decide not to undertake pooling for meta-analysis (due in part to heterogeneity), we will provide the rationale for doing so in full.

**Assessment of reporting biases**

In the case of sufficient included studies, we will examine the potential for publication bias according to the recommendations outlined in Chapter 13 of the Cochrane Handbook for Systematic Reviews of Interventions (Higgins 2021). Namely, we will attempt to characterise missingness using a funnel plot and contour-enhanced funnel plot, and discuss any observations regarding asymmetry of the results.

**Data synthesis**

We will analyse data using Review Manager Web (RevMan Web 2019). We will illustrate results using forest plots for each outcome. We will assess heterogeneity (Chi^2 P value > 0.05 and I^2 < 10%) based on the Mantel-Haenszel approach. Irrespective of the degree of heterogeneity, we will use a random-effects model to provide the most conservative estimate of the effect. This will be complemented by a mixed-effects model, in which the decision to assign a fixed or random effect to a covariate will be clinically rationalised (based on consensus decision-making).

**Subgroup analysis and investigation of heterogeneity**

We will undertake the following subgroup analyses, in part to determine whether they might explain any identified heterogeneity, but also to demonstrate the robustness of any conclusions that can be drawn from the meta-analysis of the studies.

- Follow-up time points: the primary outcome is defined as any time point more than six months from recruitment or the end of the induction phase; however, the longer the follow-up period, the more likely it is that there will be attrition of participants, thus any difference in timing of the study endpoints may explain potential heterogeneity. As such, we will assess studies in subgroups to determine if any such issues are discernable.
- Age-related differences: as no age restrictions have been applied in the inclusion/exclusion criteria, where possible, we will investigate for evidence of heterogeneity introduced by studies that combined paediatric and adult populations instead of treating them separately.
- Treatment strategy: split-dosing versus simple increases and decreases in dose are fundamentally different approaches. Moreover, studies may use different set changes in dose (e.g. in the UK, the heuristic is 25%). As such, we will undertake both a comparison of split versus simple increases and decreases, and comparisons based on the size of the suggested change in dose where data are sufficient.
- Geography: variation in baseline standard of care and approaches to managing IBD are to be expected. As such, we will undertake a comparison of studies stratified by the country in which they took place. Multicountry studies will be treated as different to single-country studies for this analysis, unless the...
vast majority of participants (> 70%) were recruited from one site.

Sensitivity analysis
We will perform sensitivity analysis to examine the effects of the following on the results of the analysis.

- Definitions of disease state: as noted previously, the definition of maintaining remission is not always the same as that of successfully achieving remission in an induction phase of a study. As such, we will carry out a sensitivity analysis to compare results based on different definitions of maintenance. We expect comparisons will be made between studies that use global clinician ratings of remission, outcome measures that combine clinician- and patient-reported items, and endoscopic definitions.
- Incomplete reporting: we will undertake an assessment of the sensitivity of the review’s conclusions to the inclusion/exclusion of studies reported only in the form of an abstract or conference proceedings.
- Unexplained loss to follow-up: for those studies in which a substantial proportion of data are missing or unavailable (i.e. > 20% of all participants recruited at baseline), we will carry out a series of sensitivity analyses using the worst case (missing = treatment failure) and best case (missing = treatment success) to determine the possible impact on the results (Gamble 2005).

We may identify additional sensitivity analyses as being relevant in the course of conducting the review. These will be explicitly acknowledged as having been retrospectively added to the statistical protocol in the final review.

Summary of findings and assessment of the certainty of the evidence
We will use the GRADE approach to assess the certainty of the available evidence (Guyatt 2008; Higgins 2021), assigning a judgment of one of the following four outcome categories to each outcome assessed:

- **High certainty:** we are very confident that the true effect lies close to that of the estimate of the effect.
- **Moderate certainty:** we are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
- **Low certainty:** our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
- **Very low certainty:** we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

Evidence from randomised controlled trials starts at high certainty and is then downgraded in a step-wise fashion based on the following five criteria: overall risk of bias, consistency of effect, imprecision, indirectness, and publication bias. We will determine the final GRADE judgements for each of the primary outcomes by consensus amongst all five review authors:

- efficacy of thiopurine-guided azathioprine treatment for maintaining clinical or endoscopic remission relative to weight-based dosing;
- the safety of the thiopurine-guided azathioprine treatment strategy relative to weight-based dosing.

**ACKNOWLEDGEMENTS**

The authors would like to thank the following editors and peer referees who provided feedback to improve the protocol: Paul Moayyedi (Co-ordinating Editor), Teo Quay (Managing Editor), Yuhong Yuan (Information Specialist), Azhar Ansari, Nanne de Boer, Lissy de Ridder (Peer Reviewers), and Lisa Winer (Copy Editor).
6-thioguanine nucleotide monitoring in azathioprine and mercaptopurine monotherapy for the treatment of inflammatory bowel disease (Protocol)

REFERENCES

Additional references

Abbass 2019

Akobeng 2003

Akobeng 2016

Alatab 2020

Battat 2017

Baumgart 2008

Benkov 2013

Best 1976

Bickston 2014

Calne 1962

Cargnin 2018

Chande 2015

Chande 2016

Colombel 2000

Cournard 2016

Cuffari 2000

Cuffari 2006
6-thioguanine nucleotide monitoring in azathioprine and mercaptopurine monotherapy for the treatment of inflammatory bowel disease

(Protocol)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.
m immunosuppression in nonimmune cells: new target for an old drug. *Journal of Immunology* 2014;192(9):4370-8. [DOI: 10.4049/jimmunol.1302527]

Marinković 2014b
Marinković G, Hamers AJ, de Vries CJM, de Waard V. 6-mercaptopurine reduces macrophage activation and gut epithelium proliferation through inhibition of GTPase Rac1. *Inflammatory Bowel Diseases* 2014;20(9):1487-95. [DOI: 10.1097/MIB.0000000000000122]

Marshall 2010

Marshall 2012

Moher 2009

Moreau 2014

Moriyama 2016

Nelson 2018

Osterman 2006

Paine 2014
6-thioguanine nucleotide monitoring in azathioprine and mercaptopurine monotherapy for the treatment of inflammatory bowel disease (Protocol)

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.
Appendix 1. CENTRAL search strategy

1. exp Inflammatory bowel diseases/
2. (inflammatory bowel disease* or IBD).tw,kw.
3. crohn*.tw,kw.
4. (colitis or regional enteritis or proctocolitis or colorectitis).tw,kw.
5. (UC or CD).tw,kw.
6. or/1-5
7. exp azathioprine/
8. azathioprin*.mp.


Walmsley 1998


Wang 2015


Wang 2016a


Wang 2016b


Warner 2018


Weinshilboum 1980


Wright 2004

9. exp Mercaptopurine/
10. mercaptopurin*.mp.
11. (Imuran or Purinethol or AZA or 6-MP or 6MP).tw,kw.
12. thiopurin*.tw,kw.
13. exp Antimetabolites/
14. (antimetabolit* or anti-metabolit*).mp.
15. or/7-14
16. 6 and 15
17. (thioguanine nucleotide or 6-TGN or 6TGN or Thiopurine methyltransferase).tw,kw.
18. (threshold* or monitor*).tw,kw.
19. exp Drug Monitoring/
20. (dose or dosing) adj3 (adjust* or strateg*).tw,kw.
21. or/17-20
22. 16 and 21

Appendix 2. MEDLINE (Ovid) search strategy
1. exp Inflammatory bowel diseases/
2. (inflammatory bowel disease* or IBD).tw,kw.
3. crohn*.tw,kw.
4. (colitis or regional enteritis or proctocolitis or colorectitis).tw,kw.
5. (UC or CD).tw,kw.
6. or/1-5
7. exp azathioprine/
8. azathioprin*.mp.
9. exp Mercaptopurine/
10. mercaptopurin*.mp.
11. (Imuran or Purinethol or AZA or 6-MP or 6MP).tw,kw.
12. thiopurin*.tw,kw.
13. exp Antimetabolites/
14. (antimetabolit* or anti-metabolit*).mp.
15. or/7-14
16. 6 and 15
17. (thioguanine nucleotide or 6-TGN or 6TGN or Thiopurine methyltransferase).tw,kw.
18. (threshold* or monitor*).tw,kw.
19. exp Drug Monitoring/
20. (dose or dosing) adj3 (adjust* or strateg*).tw,kw.
21. or/17-20
22. 16 and 21
23. randomized controlled trial.pt.
24. controlled clinical trial.pt.
25. random*.ab.
26. trial.ab.
27. groups.ab.
28. drug therapy.fs.
29. placebo.ab.
30. or/23-29
31. exp animals/ not humans.sh.
32. 30 not 31
33. 22 and 32

Lines 23-32. RCT filter: Cochrane Highly Sensitive Search Strategy for identifying randomized trials in MEDLINE: sensitivity-maximizing version (2008 revision); Ovid format. We made the following minor revision: we used “random*” instead of “randomized.ab” or “randomly.ab” to capture word variations such as “randomised, randomization, random”.

6-thioguanine nucleotide monitoring in azathioprine and mercaptopurine monotherapy for the treatment of inflammatory bowel disease (Protocol)
Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.
Appendix 3. Embase (Ovid) search strategy

1. exp inflammatory bowel disease/
2. (inflammatory bowel disease* or IBD).tw,kw.
3. crohn*.tw,kw.
4. (colitis or regional enteritis or proctocolitis or colorectitis).tw,kw.
5. (UC or CD).tw,kw.
6. or/1-5
7. exp azathioprine derivative/ or exp azathioprine/
8. azathioprin*.mp.
9. exp mercaptopurine/ or exp 6 mercaptopurine derivative/
10. mercaptopurin*.mp.
11. (Imuran or Purinethol or AZA or 6-MP or 6MP).tw,kw.
12. thiopurin*.tw,kw.
13. exp antimetabolite/
14. (antimetabolit* or anti-metabolit*).mp.
15. or/7-14
16.6 and 15
17. (thioguanine nucleotide or 6-TGN or 6TGN or Thiopurine methyltransferase).tw,kw.
18. (threshold* or monitor*).tw,kw.
19. exp drug monitoring/
20. (dose or dosing) adj3 (adjust* or strateg*).tw,kw.
21. or/17-20
22.16 and 21
23. random:.tw.
24. placebo:.mp.
25. double-blind:.tw.
26. or/23-25
27. exp animal/ not human.sh.
28.26 not 27
29.22 and 28

Lines 23-26. RCT filter. Two or more terms min difference version (https://hiru.mcmaster.ca/hiru/hedges/All-EMBASE.htm)

CONTRIBUTIONS OF AUTHORS

BAM conceived and designed the protocol, with input from MG, MP, and BH. BAM wrote the first draft of the protocol. The search strategy was designed by MG and BAM in partnership with the Cochrane Gut Review Group Information Specialist. All authors contributed to the critical review and drafting of the final manuscript. BH secured the funding specific to facilitating this project.

The corresponding author (BAM) and senior author (BH) are the joint guarantors for the protocol, and had final responsibility for the decision to submit for publication.

DECLARATIONS OF INTEREST

Bilal A Mateen: none relevant to the contents of this review.
Mehul Patel: none relevant to the contents of this review.
Anthony K Akobeng: none relevant to the contents of this review.
Morris Gordon: received travel fees to attend international scientific and training meetings such as DDW, Advances in IBD, ESPGHAN, BSPGHAN, and Cochrane-focused international events from companies including Abbott, Nutricia, Biogaia, Ferring, Allergan, and Tillots.
Bu’Hussain Hayee: none relevant to the contents of this review.
SOURCES OF SUPPORT

Internal sources

- Author BM Funding, UK
  
  Funding covering bench fees and tuition has been provided by the Kings College Hospital (Department of Gastroenterology Research Fund; Code 3339).

External sources

- Author MG Funding, UK
  
  Partial funding for MG was provided through an NIHR Cochrane Programme Grant (16/114/13) in the UK.
- Cochrane IBD Group Funding, Canada
  
  Funding for the Cochrane IBD Group (1 May 2017 to 30 April 2022) has been provided by Crohn’s and Colitis Canada (CCC).