Appendix S1

METHODS

Study site and population

A pilot, assessor-blinded, RCT was conducted in St. Mary's Hospital, Manchester, a large tertiary hospital in North West England which has more than 8,000 births annually. Babies were included if they were born to mothers carrying singleton pregnancies and booked for care at St. Mary's Hospital, were full term (37 weeks gestation or more), were in good health (as determined by the investigator) and were less than 48 h old if recruited between September 2013 and February 2014 inclusive or less than 72 h old if recruited between March 2014 and June 2014 inclusive. Mothers were excluded if they were 16 years of age or less or did not have capacity to consent. Babies were excluded if they had been admitted to Special Care Baby Unit, were having phototherapy treatment, were in another clinical trial, had any medical history preventing their participation to endpoint, had limb defects, non-traumatic impairment of epidermal integrity or evidence of skin disorder at first assessment. Normal neonatal skin variations such as erythema neonatorum/toxicum and milia were not considered to be skin disorders for this study. We set a target sample size of 100 babies to allow for 30 per group after a 10% anticipated loss to follow-up. We included those with and without a family history of atopic eczema (AE), and used stratification to ensure that cases were evenly distributed across groups. The sample size was considered to be sufficient to explore differences in outcomes and provide data capable of determining feasibility for a definitive trial (28). The trial was approved by Greater Manchester East Research Ethics Committee (13/NW/0512).

Recruitment and randomization

Women who were potentially eligible for their baby to take part were given summary study information antenatally at 28 weeks gestation. Willing participants completed a response slip in order to provide consent for the investigator to approach them after childbirth. Alternatively, women were screened postnatally by the investigator and approached following permission from the clinical team. All eligible women who gave permission to be approached following birth were provided with full study information and a verbal explanation; they were then given time to consider taking part.

Babies of women who gave consent were randomized to one of the intervention groups or the control group within 72 h of birth. Randomization was 1:1:1 via a central telephone-based service provided by The Christie Hospital NHS Foundation Trust Clinical Trials Unit. The randomization sequence was computer generated. Randomization was stratified according to whether or not there was a family history of AE, where at least one of father, mother, or sibling had a medical diagnosis of AE and had been prescribed topical steroid treatment. The randomization was in blocks within eczema history strata (yes, no) and the block size varied at random between 6 and 15 (i.e. 6, 9, 12 or 15) to guard against predictability. Allocation was concealed from the participant and independent research midwife until the point of allocation. Babies were randomized to one of 3 groups: olive oil, sunflower oil or no oil (control). Following randomization women were given the appropriate advice and materials by the independent research midwife to maintain investigator blinding. The study was assessor-blinded, and participants in the intervention groups were blinded to which oil they were using; oils were labelled X and Y. Identification and labelling of oil X and oil Y was conducted by an independent researcher at The University of Manchester, confirmed by a second independent researcher. The information was sealed in two envelopes which were kept by two independent university staff until after data analysis. Participant blinding was impossible for the control group as there is no control oil that we could be confident was safe to apply and would have no effect on skin barrier function (29).

Intervention

Olive oil and sunflower oil of specific defined formulation (William Hodgson and Co, Congleton, United Kingdom; see Table I) were provided for the intervention groups as appropriate. Oil was provided in opaque plastic dropper bottles. The first application was demonstrated by an independent research midwife who had been instructed to provide the appropriate advice. Parents then began using the oil as instructed from the day after the initial assessment. Parents applied 4 drops of oil to their baby's left forearm, left thigh and abdomen, twice a day, up until the night before their follow-up assessment at approximately 4 weeks, using a clean hand to spread the oil evenly across the treatment area. A diagrammatic laminate was provided to ensure parents applied oil to the correct areas of their baby's skin at each application. No oils were applied on the day of assessment to avoid any interference with results that may have been caused by oil residues, and to maintain assessor blinding. Parents were asked to return any unused oil to a box when they attended for follow-up assessment, prior to meeting the investigator. They were asked not to discuss their treatment allocation with the investigator at any opportunity. Parents in all 3 groups were asked not to use any other skincare products on the 3 study sites; water only was advocated. Parents received a weekly phone call to ask about any product use, health professional consultations, medication prescriptions and whether there were any rashes or skin concerns. These data were recorded for compliance and safety purposes.

Assessment of trial outcomes

All measurements were taken by the investigator who remained blind to the treatment allocation. It was intended that all measurements would take place in the same clinical room to maintain a controlled environment. In the final phase of the study, home visits were offered to those who advised that they were unable to return to the hospital for follow-up. Home visits were offered due to the higher than anticipated loss to follow-up, and to assess their feasibility for a definitive trial. Data were collected at two time points. The first assessment was conducted at baseline prior to discharge from the hospital. A second assessment was made at 4 weeks ± 5 days. Measurement consistency was achieved for each study site by measuring from anatomical markers such as the skin crease of the wrist to midpoint on the volar forearm, above the patella to midpoint on thigh, and above umbilicus to midpoint to nipple line for upper abdomen. For the primary outcome measures, two measurements were taken at each treatment site. In between measurements, 3 consecutive D-Squame discs (CuDerm Corporation, Dallas, TX, USA) were applied to and removed from the site. The D-Squame discs remove (tape-stripping) the very top skin cells, which are already dead and about to be lost naturally from the surface of the skin in a process known as desquamation. To ensure that tape stripping did not occur in the same site at follow-up as baseline assessment, measurements were conducted on the skin just below midpoint at baseline and just above midpoint at follow-up.

Primary outcomes

ATR-FTIR spectroscopy. The change in structure of the lipid lamellae, a determinant of stratum corneum (SC) permeability barrier function (30), was assessed between 48 h and 4 weeks

following birth using ATR-FTIR spectroscopy. This technique has been used previously to demonstrate the effect of oleic acid on skin barrier (8). ATR-FTIR spectra were collected non-invasively using a silver halide fibre-optic probe (FTIR Flexispec PIR 900, Art Photonics, Berlin, Germany) attached to a Nicolet iS50 FTIR Spectrometer (Thermo Fisher Scientific Inc., Waltham, USA), equipped with a cooled mercury-cadmium-telluride detector and purged with dry nitrogen. We collected a mean of 40 scans per measurement (resolution 4 wavenumbers). At each spectroscopy measurement site on the skin surface an absorbance spectrum was collected on intact skin and following the application and removal of 3 consecutive D-Squame discs (tape-stripping) to reassess the deeper corneocyte layers of the SC. Data analysis of absorbance spectra was performed in Omnic 9.0 and TQuant (Thermo Fisher Scientific Inc., Waltham, USA). The difference in the quantity of lipids and lipid esters in the skin was determined based upon the change in peak intensities of the spectral regions centred on ~2,920 and ~2,850 wavenumbers, corresponding to the symmetric and asymmetric stretching of the CH2 group of all lipids, and 1,740 wavenumbers, corresponding to lipid esters of triglycerides in sebum and topically applied oils, respectively (31). Quantities were normalized to regions of the spectra showing no absorbance, at 3,800 and 1,800 wavenumbers respectively, to account for differences in contact pressure between the skin and the probe. Lipid chain conformation ($\nu_{asym}CH_2$ COG) was based on the location (centre of gravity: COG) of the peak between ~2,853 and ~2,848 wavenumbers, corresponding to the asymmetric stretching of the CH2 bond of lipids (32, 33). A peak centre of gravity at ~2,848 wavenumbers corresponds to tightly packed lipid chains, and is associated with optimum skin barrier function, whereas higher wavenumbers indicate increasing lipid fluidity and decreasing skin barrier function. Lateral chain packing was determined from the second derivative reflectance spectra by measuring the full width at half maximum (FWHM) of the spectral region centred at 1,468 wavenumbers (30). A change in the width of this region corresponds to changes in lateral lipid chain packing. Highly ordered orthorhombic packing of lipids is indicated by a FWHM of ≥ 11 wavenumbers. A higher proportion of orthorhombic structuring throughout the depth of the SC is associated with improved skin barrier function. The difference in the quantity of surfactants in the skin, measured to assess adherence regarding use of wash products, was determined based upon the change in peak intensity of the spectral region centred on 1,240 wavenumbers, corresponding to the sulphur group of surfactants found in wash products, normalized to the reference region at 1,800 wavenumbers (34).

Trans-epidermal water loss (TEWL). This outcome measured the rate of change of basal trans-epidermal water loss (TEWL) between 48 h and 4 weeks after birth. TEWL, a validated measure of skin barrier function (35), is defined as the flux of water vapour evaporating from the skin surface, and was measured using a closed chamber TEWL instrument (Biox Aquaflux Model AF200). The lead investigator took the measurements at both time points, in accord with published guidelines for TEWL measurements (36). Measurements were taken at each study site twice, before and after tape stripping.

Secondary outcomes

Stratum corneum hydration and skin surface pH. The change in SC hydration and skin surface pH between 48 h and 4 weeks were measured at the same times and sites as the primary outcome measures using a Corneometer® Model CM825 [Courage & Khazaka electronic GmbH, Köln, Germany] and skin pH meter® Model PH905 [Courage & Khazaka electronic GmbH].

Clinical observations. Changes in the skin were observed and recorded by the investigator at baseline and follow up (erythema, dryness and scaling, need for medical products/attention) between 48 h and 4 weeks. The investigator assessed the babies' skin according to a modified Neonatal Skin Condition Score (NSCS; 37). Rating was based on severity of dryness and scaling. A score of zero indicated no evidence of abnormal skin increasing to a score of 4 which indicated a degree of severity. Details of skin condition and need for medical products/attention were also collected via a weekly telephone questionnaire conducted with mothers by the investigator. Erythema was measured using a Mexameter® Model MX18 probe at each visit [Courage & Khazaka electronic GmbH].

Analysis

Data were double-entered into IBM SPSS Statistics version 20 and analysed in version 22, with the two data files cross-checked for errors. In accordance with recommended practice for pilot studies (28), the main analyses were descriptive, involving the estimation of recruitment rates, attrition rates, adherence rates, means and standard deviations of primary and secondary outcomes by group at baseline and 4 weeks, and 95% confidence intervals for differences of means of change scores of primary and secondary outcomes between groups at 4 weeks. Missing values at 4 weeks were not carried forward or imputed; descriptive analysis at 4 weeks was based on complete data, compared by randomization group. The latter comparisons were confirmed by analysis of covariance.